Towards a non-local density functiona description of exchange and correlation

Rushton, Philip Peter

University of Durham

Towards a Non-Local Density Functional Description of Exchange and Correlation

Philip Peter Rushton

Departments of Chemistry and Physics
University of Durham

November 2002

A thesis submitted for the degree of
Doctor of Philosophy

Abstract

Kohn-Sham density functional theory (DFT) is a popular method used to investigate the properties of matter. Although exact in principle, DFT is in practice limited by a single approximation for the exchange-correlation functional - the quantity that describes the many-body interactions between electrons. This thesis is concerned with developing improved exchange-correlation functionals for use in practical DFT calculations. The standard functional currently used in solid state physics, and also popular in quantum chemistry, is the generalised gradient approximation (GGA), which requires only the local density, $n(\mathbf{r})$, and the density gradient, $\nabla n(\mathbf{r})$, as input. A flexible semi-empirical GGA form, containing 15 free parameters that are fitted to near-exact molecular data is implemented within the plane-wave pseudopotential (PW-PP) Kohn-Sham scheme, to assess the possibility of employing semi-empirical GGAs in solid state applications. Self-consistent calculations performed for several bulk semiconductor properties using this GGA reveal that, overall, no improvement is attained over a conventional non-empirical GGA used in solid state physics. The remainder of the thesis focuses on a fully non-local functional known as the weighted density approximation (WDA), which utilises the global density of a system, $n(\mathbf{r}')$, as input. An efficient computational algorithm is devised for use within the PW-PP formalism which enables fully self-consistent WDA calculations to be performed. Physical properties are shown to be intimately related to the particular form used for the pair-correlation function, $g_{\mathrm{XC}}^{\mathrm{WDA}}(\mathbf{r},\mathbf{r}')$, and by comparing with recent variational Monte Carlo (VMC) data, it is shown that the forms that provide a good description of exchange-correlation holes, $n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')$, also lead to the most accurate bulk properties. For strongly inhomogeneous electron gas systems, the WDA provides close agreement with the VMC method for a variety of exchange-correlation quantities. The success of the fully non-local approach given by the WDA for other model electron gas systems studied suggests that the WDA is a very promising functional.

Declaration

The work presented in this thesis was carried out in the Department of Chemistry and the Condensed Matter Theory Group in the Department of Physics, University of Durham, between October 1999 and October 2002. The work has not been submitted for any other degree in the University of Durham or elsewhere, and unless stated otherwise, is the original work of the author.

Copyright

Philip P. Rushton

November 2002

Acknowledgements

I would like to thank my supervisors Dave Tozer and Stewart Clark for assistance and valuable discussions over the past three years. Many thanks go to Lydia Heck for solving my computer problems and to Nigel Metcalf for installing Linux on my laptop, which has proved invaluable in the writing of this thesis. I am very grateful to Randy Hood and Tony Cancio for supplying computer codes and data for the generation of the VMC exchange-correlation holes in Si, and also to Maziar Nekovee for useful discussions regarding the VMC work on the cosine-wave electron gas.

Outside work, I thank everyone who has been involved with Durham University Shotokan Karate Club (DUSKC) over the past seven years for providing the entertainment and adventures. I especially thank Steve whom I have shared many funny moments. Finally, I am indebted to my family for their constant support and kindness throughout the entirety of my education, and to Penny, for always being there for me.

For Mum, Dad, Stewart, Kate, and Penny

Contents

Chapter 1

The Quantum Many-Body Problem

1.1 Introduction

A significant part of condensed matter physics and chemistry would be solved if the electronic structure of atoms, molecules and solids could be determined exactly. This however is a formidable task for two main reasons. Firstly, electrons in matter must be treated using the laws of quantum mechanics rather than classical physics - the quantum length scale is set by Planck's constant h, and the onset of quantum effects occurs when the de Broglie wavelength of a particle, λ, given by

$$\lambda = \frac{h}{p} \tag{1.1}$$

is comparable to the average inter-particle separation. Rearranging the energy-momentum equation $E = p^2/2m_e$, and the electronic thermal energy relationship $E \sim k_B T$, leads to a relation for the de Broglie wavelength of an electron given in terms of the electron mass m_e and temperature T,

$$\lambda = \frac{h}{\sqrt{2m_e k_B T}}. \tag{1.2}$$

For solid-state systems the average inter-electron separation is usually represented by the Seitz radius, r_s, which is the radius of a sphere whose volume encloses a single electron in the system, and is normally written in terms of the Bohr radius, $a_0 \sim 0.529 \times 10^{-10}$m,

$$\frac{r_s}{a_0} = \left(\frac{3}{4\pi n_0}\right)^{1/3}, \tag{1.3}$$

where n_0 is the average electron density. For most systems of interest r_s typically ranges from 0.1 up to 10, which means that the electron de Broglie wavelength is larger or comparable to the average separation up to $T \sim 10^4 K$, according to relation (1.2). Therefore within this temperature range the de Broglie wavelength of the electrons overlap and the interactions between the electrons become quantum-mechanically correlated.

The second problematic issue concerns the number of electrons that are involved - the coupling of the electron interactions due to de Broglie wavelength overlap renders an analytic solution impossible for systems with more than one electron, and the complexity grows dramatically with increasing electron number. It is for these reasons that the electronic structure of matter is known as the quantum many-body problem.

The quantum many-body problem is unusual within the realm of theoretical physics because the equations required for an exact solution are known. The properties of any (non-relativistic) time-independent quantum system can be determined by solving the Schrödinger equation [1],

$$\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_N), \tag{1.4}$$

where $\hat{H}$, $\Psi(\mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_N)$ and E are the Hamiltonian, many-body wavefunction and total energy of the system. Matter consists of electrons and nuclei interacting with each other Coulombically, consequently the Hamiltonian for any such system is given by,

$$\hat{H} = -\sum_{i=1}^{M} \frac{\hbar^2}{2m_{z_i}} \nabla^2_{\mathbf{R}_i} - \sum_{i=1}^{N} \frac{\hbar^2}{2m_e} \nabla^2_{\mathbf{r}_i} + \frac{1}{4\pi\epsilon_0} \sum_{i}^{M} \sum_{j>i}^{M} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}$$

$$- \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{Z_j e}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (1.5)$$

where M and N are the number of nuclei and electrons in the system, m_z, Z and $\mathbf{R}$ are the mass, charge and position of the nuclei, m_e and e are the mass and charge of an electron, and $\mathbf{r}$ represents the position of the electrons. The first two terms in (1.5) are the kinetic energy contributions from the nuclei and the electrons respectively, and the rest are Coulombic potential energy terms arising from the ion-ion repulsion, ion-electron attraction and the electron-electron repulsion respectively. Although in principle everything is known exactly, the Schrödinger equation (1.4) with this Hamiltonian is simply too difficult to solve directly. Hence, the quantum many-body problem is centred upon finding intelligent approximations to the Hamiltonian (1.5) and the many body wavefunction Ψ, that retain the correct physics and are computationally tractable to solve.

The first simplification of this problem is attributed to Born and Oppenheimer [2] who recognised that in most cases the nuclear and electronic degrees of freedom can be decoupled since they exhibit vastly different dynamics - the nuclei are of order $\sim 10^3$ times heavier than the electrons and so are considered to be stationary with respect to the electrons. The electrons therefore move within a fixed external potential due to the nuclei. Within the Born-Oppenheimer approximation the complexity of the full many-body Hamiltonian (1.5) reduces to that of an electronic Hamiltonian,

$$\hat{H} = - \sum_{i=1}^{N} \frac{\hbar^2}{2m_e} \nabla_i^2 - \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{Z_j e}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.6)$$

Solving the Schrödinger equation with the above Hamiltonian is however still too complex for most cases since the many-electron wavefunction contains $3N$ variables, which for a solid containing $N \sim 10^{26}$ electrons, is simply an intractable number of degrees of freedom.

Devising accurate schemes to approximate the many-electron problem has been an important goal since the founding of quantum mechanics in

the early 1900s. Several notable advances have been made, starting from Thomas-Fermi theory in the late 1920s [3, 4] which made a significant conceptual presumption by having the electron density, $n(\mathbf{r})$, as the central unknown variable, rather than the many-electron wavefunction. This approach simplified the problem considerably since the density contains just three degrees of freedom, namely the x, y, z coordinates of the system. In 1930 came Hartree-Fock theory [5, 6] which builds upon the single-particle approximation proposed earlier by Hartree [7], but in addition correctly accounts for the exchange interactions between electrons that are a consequence of the Pauli principle, by antisymmetrising the single-particle functions $\psi_i(\mathbf{r}_i s_i)$,

$$\Psi(\mathbf{r}_1 s_1, \mathbf{r}_2 s_2 \ldots \mathbf{r}_N s_N) \approx \frac{1}{\sqrt{N}} \mathcal{A} \left[\psi_1(\mathbf{r}_1 s_2)\, \psi_2(\mathbf{r}_2 s_2) \ldots \psi_N(\mathbf{r}_N s_N) \right] . \quad (1.7)$$

The symbol $\mathcal{A}$ represents the antisymmetric nature of the single-particle products, and s_i gives the spin dependence. This has the desired effect of decoupling the $3N$ degrees of freedom in the many-electron wavefunction, and so allows each degree of freedom to be solved independently.

A significant leap in electronic structure theory was made in 1964 with the remarkable theorems of density functional theory (DFT), proved by Hohenberg and Kohn [8]. DFT allows the ground-state properties of a many-electron system to be determined exactly through the electron density $n(\mathbf{r})$, and therefore in a computationally tractable manner, however DFT is only a proof of existence, it does not give details of how this can be achieved in practice. In 1965 Kohn and Sham [9] devised an ingeniously practical single-particle scheme for performing DFT calculations, which is still exact, in principle. The price to be paid for the benefits of Kohn-Sham DFT is that the single-particle Hamiltonian is only partly known in practice - approximations must be made for a single unknown component that accounts for electron many-body effects, known as exchange and correlation. Improving the exchange-correlation approximation in DFT is the object of this thesis.

The many-body methods just introduced will be discussed in more detail

in the following sections. Unless otherwise stated, all equations, figures and tables in the remainder of this thesis will use atomic units, whereby $\hbar = e = m_e = 4\pi\epsilon_0 = 1$.

1.2 Thomas-Fermi Theory

One of the earliest tractable schemes for solving the many-electron problem was proposed by Thomas and Fermi [3, 4]. In this model the electron density $n(\mathbf{r})$ is the central variable rather than the wavefunction, and the total energy of a system is written as a functional $E^{\mathrm{TF}}[n(\mathbf{r})]$, where square brackets are used to enclose the argument of the functional, which in this case is the density. The Thomas-Fermi energy functional is composed of three terms,

$$E^{\mathrm{TF}}[n(\mathbf{r})] = A_k \int n(\mathbf{r})^{5/3}\, d\mathbf{r} + \int n(\mathbf{r})\, v_{\mathrm{ext}}(\mathbf{r})\, d\mathbf{r} + \frac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}\, d\mathbf{r}'\,. \tag{1.8}$$

The first term is the electronic kinetic energy associated with a system of non-interacting electrons in a homogeneous electron gas. This form is obtained by integrating the kinetic energy density of a homogeneous electron gas $t_0[n(\mathbf{r})]$ [10, 11],

$$T^{\mathrm{TF}}[n(\mathbf{r})] = \int t_0[n(\mathbf{r})]\, d\mathbf{r}\,, \tag{1.9}$$

where $t_0[n(\mathbf{r})]$ is obtained by summing all of the free-electron energy states $\varepsilon = k^2/2$, up to the Fermi wavevector $k_{\mathrm{F}} = [3\pi^2 n(\mathbf{r})]^{1/3}$,

$$t_0[n(\mathbf{r})] = \frac{2}{(2\pi)^3}\int \frac{k^2}{2}\, n_k\, d\mathbf{k} = \frac{1}{2\pi^2}\int_0^{k_{\mathrm{F}}} k^4\, dk\,, \tag{1.10}$$

n_k is the density of allowed states in reciprocal-space. This leads to the form given in (1.8) with coefficient $A_k = \frac{3}{10}(3\pi^2)^{2/3}$. The power-law dependence on the density can also be established on dimensional grounds [12]. The second term is the classical electrostatic energy of attraction between the nuclei and

the electrons, where $v_{\text{ext}}(\mathbf{r})$ is the static Coulomb potential arising from the nuclei,

$$v_{\text{ext}}(\mathbf{r}) = -\sum_{j=1}^{M} \frac{Z_j}{|\,\mathbf{r} - \mathbf{R}_j\,|}\,. \tag{1.11}$$

Finally, the third term in (1.8) represents the electron-electron interactions of the system, and in this case is approximated by the classical Coulomb repulsion between electrons, known as the Hartree energy.

To obtain the groundstate density and energy of a system, the Thomas-Fermi equation (1.8) must be minimised subject to the constraint that the number of electrons is conserved. This type of constrained minimisation problem, which occurs frequently within many-body methods, can be performed using the technique of Lagrange multipliers. In general terms, the minimisation of a functional $F[f]$, subject to the constraint $C[f]$, leads to the following stationary condition,

$$\delta\Big(F[f] - \mu C[f]\Big) = 0\,, \tag{1.12}$$

where μ is a constant known as the Lagrange multiplier. Minimising (1.12) leads to the solution of the corresponding Euler equation,

$$\frac{\delta F[f]}{\delta f} - \mu \frac{\delta C[f]}{\delta f} = 0\,. \tag{1.13}$$

Applying this method to (1.8) leads to the stationary condition,

$$\delta\Big\{E^{\text{TF}}[n(\mathbf{r})] - \mu\left(\int n(\mathbf{r})\,d\mathbf{r} - N\right)\Big\} = 0\,, \tag{1.14}$$

which yields the so-called Thomas-Fermi equations,

$$\frac{5}{3}A_k\, n(\mathbf{r})^{2/3} + v_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\,\mathbf{r} - \mathbf{r}'\,|}\,d\mathbf{r}' - \mu = 0\,, \tag{1.15}$$

that can be solved directly to obtain the groundstate density,

Thomas-Fermi theory suffers from many deficiencies, probably the most serious defect is that it does not predict bonding between atoms [13, 14, 15], so molecules and solids cannot form in this theory. The main source of

error comes from approximating the kinetic energy in such a crude way. The kinetic energy represents a substantial portion of the total energy of a system and so even small errors prove disastrous. Another shortcoming is the over-simplified description of the electron-electron interactions, which are treated classically and so do not take account of quantum phenomenon such as the exchange interaction.

1.3 Dirac Exchange

Shortly after the introduction of Thomas-Fermi theory, Dirac [16] developed an approximation for the exchange interaction based on the homogeneous electron gas. The resulting formula is simple, and is also a local functional of the density,

$$E_{\mathrm{x}}[n(\mathbf{r})] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int n(\mathbf{r})^{4/3}\,d\mathbf{r}\,. \tag{1.16}$$

Relation (1.16) is usually written in terms of the exchange energy density $\varepsilon_{\mathrm{x}}[n(\mathbf{r})]$ as,

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n(\mathbf{r})] = \int n(\mathbf{r})\,\varepsilon_{\mathrm{xc}}[n(\mathbf{r})]\,d\mathbf{r}\,, \tag{1.17}$$

where $\varepsilon_{\mathrm{xc}}[n(\mathbf{r})]$ can be given simply in terms of the Seitz radius r_s,

$$\varepsilon_{\mathrm{x}}[n(\mathbf{r})] \;=\; -\frac{3}{4}\left(\frac{9}{4\pi^2}\right)^{1/3}\frac{1}{r_s} \;\approx\; -\frac{0.4582}{r_s}\,. \tag{1.18}$$

The Dirac exchange term was naturally incorporated into to Thomas-Fermi theory by simply adding (1.16) to (1.8), and including the term, $4/3\,\varepsilon[n(\mathbf{r})]$, in the corresponding Euler-Lagrange equation (1.15). The inclusion of local exchange did not improve the Thomas-Fermi method [6].

1.4 Hartree-Fock Theory

The simplest way to approximate electron-electron interactions is through the Hartree approximation, where the true N-electron wavefunction Ψ is

replaced by a product of single-particle orbitals, $\psi_i(\mathbf{r}_i s_i)$,

$$\Psi(\mathbf{r}_1 s_1, \mathbf{r}_2 s_2, \ldots, \mathbf{r}_N s_N) = \frac{1}{\sqrt{N}} \psi_1(\mathbf{r}_1 s_1)\, \psi_2(\mathbf{r}_2 s_2) \ldots \psi_N(\mathbf{r}_N s_N)\,, \quad (1.19)$$

where $\psi_i(\mathbf{r}_i s_i)$ is composed of a spatial function $\phi_i(\mathbf{r}_i)$, and an electron spin function $\sigma(s_i)$ such that,

$$\psi_i(\mathbf{r}_i) = \phi_i(\mathbf{r}_i)\, \sigma(s_i)\,, \quad (1.20)$$

and $\sigma = \alpha, \beta$ represent up-spin and down-spin electrons respectively. However, as mentioned previously, the Hartree approximation does not account for exchange interactions since (1.19) does not satisfy,

$$\Psi(\mathbf{r}_1 s_1, \ldots, \mathbf{r}_i s_i, \ldots, \mathbf{r}_j s_j, \ldots \mathbf{r}_N s_N) = -\Psi(\mathbf{r}_1 s_1, \ldots, \mathbf{r}_j s_j, \ldots, \mathbf{r}_i s_i, \ldots \mathbf{r}_N s_N)\,, \quad (1.21)$$

under the interchange of particle coordinates, which is required by the exclusion principle.

This problem was rectified by the Hartree-Fock approximation [5] which accounts for electron exchange interactions by writing the wavefunction as an antisymmetrised product of orbitals. The Hartree-Fock wavefunction Ψ_{HF} amounts to a linear combination of the terms in (1.19), which includes all permutations of the electron coordinates with the corresponding weights ± 1, i.e.

$$\begin{aligned} \Psi_{\text{HF}} \;=\; & \frac{1}{\sqrt{N!}} \big[\psi_1(\mathbf{r}_1 s_1)\, \psi_2(\mathbf{r}_2 s_2) \ldots \psi_N(\mathbf{r}_N s_N) \\ & -\psi_1(\mathbf{r}_2 s_2)\, \psi_2(\mathbf{r}_1 s_1) \ldots \psi_N(\mathbf{r}_N s_N) + \ldots \big]\,, \end{aligned} \quad (1.22)$$

and so fulfils (1.21). In 1951 Slater [6] realised that the Hartree-Fock wavefunction can be efficiently represented as an $N \times N$ determinant, now known as a Slater determinant:

$$\Psi_{\text{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1 s_1) & \psi_1(\mathbf{r}_2 s_2) & \ldots & \psi_1(\mathbf{r}_N s_N) \\ \psi_2(\mathbf{r}_1 s_1) & \psi_2(\mathbf{r}_2 s_2) & \ldots & \psi_2(\mathbf{r}_N s_N) \\ \vdots & \vdots & & \vdots \\ \psi_N(\mathbf{r}_1 s_1) & \psi_N(\mathbf{r}_2 s_2) & \ldots & \psi_N(\mathbf{r}_N s_N) \end{vmatrix}\,, \quad (1.23)$$

where the orbitals are subject to the orthonormal constraint,

$$\int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\,d\mathbf{r} = \langle\psi_i|\psi_j\rangle = \delta_{ij}\,. \tag{1.24}$$

The Slater determinant can also be written in shorthand notation as,

$$\Psi_{\text{HF}} = \frac{1}{\sqrt{N!}}\det[\psi_1(\mathbf{r}_1 s_1)\psi_2(\mathbf{r}_2 s_2)\ldots\psi_N(\mathbf{r}_N s_N)]\,. \tag{1.25}$$

The Hartree-Fock energy can be evaluated by taking the expectation value of the Hamiltonian (1.6) with the above Slater determinant. This yields,

$$\begin{aligned} E_{\text{HF}} &= \langle\Psi_{\text{HF}}|\hat{H}|\Psi_{\text{HF}}\rangle \\ &= \sum_i^N \int \psi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r})\right)\psi_i(\mathbf{r})\,d\mathbf{r} \\ &\quad + \frac{1}{2}\sum_i^N\sum_j^N \int\int \frac{|\psi_i(\mathbf{r})|^2\,|\psi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}' \\ &\quad - \frac{1}{2}\sum_i^N\sum_j^N \int\int \frac{\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}')\psi_j^*(\mathbf{r}')\psi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\,\delta_{s_i s_j}\,d\mathbf{r}\,d\mathbf{r}'\,. \end{aligned} \tag{1.26}$$

The last term is of significant interest since it arises from the antisymmetric nature of the Hartree-Fock wavefunction - it vanishes when $s_i \neq s_j$, which is an artefact of the Pauli principle. Consequently this term is called the exchange energy E_{X}. It should also be noted that in practice an extra term due to the repulsion energy between the ions must be added to (1.26) in order to obtain the total energy of the system.

1.4.1 The Self-Consistent Field

The Hartree-Fock groundstate energy E_0^{HF} is obtained by minimising (1.26) with respect to the variation of the orbitals, subject to the constraint that the orbitals remain orthonormal (1.24). This is another constrained minimisation problem that can be performed using the Euler-Lagrange method. The corresponding stationary condition is given by:

$$\delta\left(E_0^{\text{HF}} - \sum_i^N\sum_j^N \varepsilon_{ij}(\langle\psi_i|\psi_j\rangle - 1)\right) = 0\,, \tag{1.27}$$

where the Lagrange multipliers, ε_{ij}, form a Hermitian matrix which can be diagonalised by a unitary transformation of the orbitals. The so-called Hartree-Fock (HF) equations (in canonical form) are therefore given by,

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + \sum_j^N \int \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}'\right)\psi_i(\mathbf{r}) - \sum_j^N \int \frac{\psi_i(\mathbf{r}')\psi_j^*(\mathbf{r}')\psi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\,\delta_{s_i s_j}\, d\mathbf{r}' = \varepsilon_i \psi_i(\mathbf{r})\,. \quad (1.28)$$

In general the Hartree-Fock equations cannot be solved analytically. One exception is for the homogeneous electron gas, where the constant external potential leads to plane wave solutions that result in the local exchange energy derived by Dirac (1.16). In other situations, the Hartree-Fock equations are solved using an iterative process known as the self-consistent field procedure. Since the desired orbitals also make up their own one-electron effective potential in (1.28), the set of orbitals $\{\psi_i(\mathbf{r})\}$ that give rise to the same set after solving (1.28) are known as the self-consistent orbitals, and they are the groundstate orbitals for that system within the Hartree-Fock approximation. The self-consistent procedure starts with an initial guess for the orbitals, and successive iterations are performed with new orbitals until the self-consistent condition is achieved.

1.4.2 Correlation

Hartree-Fock theory is not an exact theory simply because it only considers a single determinant for the electron wavefunction, and this is only a small subset of the total number of allowable wavefunctions. Consequently, it is highly unlikely that the true wavefunction is contained within this subset. The only case when a single determinant is exact is for a non-interacting system of electrons.

In real systems the motions of electrons are more correlated than the mean-field description provided by Hartree-Fock. The interaction energy

missed by Hartree-Fock is commonly termed the correlation energy E_{C} [17],

$$E_{\mathrm{C}} = E_0 - E_{\mathrm{HF}} \tag{1.29}$$

where E_0 is the exact groundstate energy. Since Hartree-Fock is a variational method, i.e. $E^{\mathrm{HF}} \geq E_0$, the correlation energy is a negative quantity according to (1.29), the exception is for a one-electron system, where in this case Hartree-Fock theory is exact and $E_{\mathrm{C}} = 0$.

A natural way to incorporate correlation effects beyond the Hartree-Fock level is to mix a linear combination of Slater determinants corresponding to excited state configurations. These post Hartree-Fock methods, such as configuration interaction, coupled-cluster and Møller-Plesset theory have been extensively developed in quantum chemistry [18], and although the approach may be systematic, the computational cost increases dramatically with excitation level. As a result, the best correlated methods are currently limited to small systems such as atoms and small molecules.

1.5 Density Functional Theory

As discussed earlier, Thomas and Fermi were the first to contemplate a model for the electron many-body problem based uniquely on the electron density $n(\mathbf{r})$. Due to the severe shortcomings of this method, they probably never imagined that an exact theory could be based on the density. However, almost forty years later, Hohenberg and Kohn proved in a seminal paper [8] that this was indeed possible. In two remarkably powerful theorems they formally established the electron density as the central quantity describing electron interactions, and so devised the formally exact groundstate method known as density functional theory (DFT). The two Hohenberg-Kohn theorems are now described.

1.5.1 The Hohenberg-Kohn Theorems

The Hohenberg-Kohn theorems relate to any system consisting of electrons moving under the influence of an external potential $v_{\text{ext}}(\mathbf{r})$. Stated simply they are as follows:

Theorem 1.

The external potential $v_{ext}(\mathbf{r})$, and hence the total energy, is a unique functional of the electron density $n(\mathbf{r})$.

The energy functional $E[n(\mathbf{r})]$ alluded to in the first Hohenberg-Kohn theorem can be written in terms of the external potential $v_{\text{ext}}(\mathbf{r})$ in the following way,

$$E[n(\mathbf{r})] = \int n(\mathbf{r})\, v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} + F[n(\mathbf{r})]\,, \tag{1.30}$$

where $F[n(\mathbf{r})]$ is an unknown, but otherwise universal functional of the electron density $n(\mathbf{r})$ only. Correspondingly, a Hamiltonian for the system can be written such that the electron wavefunction Ψ that minimises the expectation value gives the groundstate energy (1.30) (assuming a non-degenerate groundstate),

$$E[n(\mathbf{r})] = \langle \Psi | \hat{H} | \Psi \rangle\,. \tag{1.31}$$

The Hamiltonian can be written as,

$$\hat{H} = \hat{F} + \hat{V}_{\text{ext}}\,, \tag{1.32}$$

where $\hat{F}$ is the electronic Hamiltonian consisting of a kinetic energy operator $\hat{T}$ and an interaction operator $\hat{V}_{\text{ee}}$,

$$\hat{F} = \hat{T} + \hat{V}_{\text{ee}}\,. \tag{1.33}$$

The electron operator $\hat{F}$ is the same for all N-electron systems, so $\hat{H}$ is completely defined by the number of electrons N, and the external potential $v_{\text{ext}}(\mathbf{r})$.

The proof of the first theorem is remarkably simple and proceeds by *reductio ad absurdum.* Let there be two different external potentials, $v_{\text{ext},1}(\mathbf{r})$ and $v_{\text{ext},2}(\mathbf{r})$, that give rise to the same density $n_0(\mathbf{r})$. The associated Hamiltonians, $\hat{H}_1$ and $\hat{H}_2$, will therefore have different groundstate wavefunctions, Ψ_1 and Ψ_2, that each yield $n_0(\mathbf{r})$. Using the variational principle [19], together with (1.31) yields,

$$E_1^0 < \langle \Psi_2 | \hat{H}_1 | \Psi_2 \rangle = \langle \Psi_2 | \hat{H}_2 | \Psi_2 \rangle + \langle \Psi_2 | \hat{H}_1 - \hat{H}_2 | \Psi_2 \rangle \quad (1.34)$$

$$= E_2^0 + \int n_0(\mathbf{r}) [v_{\text{ext},1}(\mathbf{r}) - v_{\text{ext},2}(\mathbf{r})] \, d\mathbf{r} \quad (1.35)$$

where E_1^0 and E_2^0 are the groundstate energies of $\hat{H}_1$ and $\hat{H}_2$ respectively. It is at this point that the Hohenberg-Kohn theorems, and therefore DFT, apply rigorously to the groundstate only. An equivalent expression for (1.34) holds when the subscripts are interchanged. Therefore adding the interchanged inequality to (1.35) leads to the result:

$$E_1^0 + E_2^0 < E_2^0 + E_1^0 \quad (1.36)$$

which is a contradiction, and as a result the groundstate density uniquely determines the external potential $v_{\text{ext}}(\mathbf{r})$, to within an additive constant. Stated simply, the electrons determine the positions of the nuclei in a system, and also all groundstate electronic properties, because as mentioned earlier, $v_{\text{ext}}(\mathbf{r})$ and N completely define $\hat{H}$.

Theorem 2.

The groundstate energy can be obtained variationally: the density that minimises the total energy is the exact groundstate density.

The proof of the second theorem is also straightforward: as just shown, $n(\mathbf{r})$ determines $v_{\text{ext}}(\mathbf{r})$, N and $v_{\text{ext}}(\mathbf{r})$ determine $\hat{H}$ and therefore Ψ. This ultimately means Ψ is a functional of $n(\mathbf{r})$, and so the expectation value of

$\hat{F}$ is also a functional of $n(\mathbf{r})$, i.e.

$$F[n(\mathbf{r})] = \langle\psi|\hat{F}|\psi\rangle\,. \tag{1.37}$$

A density that is the ground-state of some external potential is known as v-representable. Following from this, a v-representable energy functional $E_v[n(\mathbf{r})]$ can be defined in which the external potential $v(\mathbf{r})$ is unrelated to another density $n'(\mathbf{r})$,

$$E_v[n(\mathbf{r})] = \int n'(\mathbf{r})\, v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} + F[n'(\mathbf{r})]\,, \tag{1.38}$$

and the variational principle asserts,

$$\langle\psi'|\hat{F}|\psi'\rangle + \langle\psi'|\hat{V}_{\text{ext}}|\psi'\rangle \;>\; \langle\psi|\hat{F}|\psi\rangle + \langle\psi|\hat{V}_{\text{ext}}|\psi\rangle \tag{1.39}$$

where ψ is the wavefunction associated with the correct groundstate $n(\mathbf{r})$. This leads to,

$$\int n'(\mathbf{r})\, v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} + F[n'(\mathbf{r})] \;>\; \int n(\mathbf{r})\, v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} + F[n(\mathbf{r})]\,, \tag{1.40}$$

and so the variational principle of the second Hohenberg-Kohn theorem is obtained,

$$E_v[n'(\mathbf{r})] \;>\; E_v[n(\mathbf{r})]\,. \tag{1.41}$$

Although the Hohenberg-Kohn theorems are extremely powerful, they do not offer a way of computing the ground-state density of a system in practice. About one year after the seminal DFT paper by Hohenberg and Kohn, Kohn and Sham [9] devised a simple method for carrying-out DFT calculations, that retains the exact nature of DFT. This method is described next.

1.5.2 The Kohn-Sham Formulation

The Kohn-Sham formulation centres on mapping the full interacting system with the real potential, onto a fictitious non-interacting system whereby

the electrons move within an effective "Kohn-Sham" single-particle potential $v_{\rm KS}(\mathbf{r})$. The Kohn-Sham method is still exact since it yields the same groundstate density as the real system, but greatly facilitates the calculation.

First consider the variational problem presented in the second Hohenberg-Kohn theorem - the groundstate energy of a many-electron system can be obtained by minimising the energy functional (1.30), subject to the constraint that the number of electrons N is conserved, which leads to,

$$\delta\left[F[n(\mathbf{r})] + \int v_{\rm ext}(\mathbf{r})n(\mathbf{r})\,d\mathbf{r} - \mu\left(\int n(\mathbf{r})\,d\mathbf{r} - N\right)\right] = 0\,, \tag{1.42}$$

and the corresponding Euler equation is given by,

$$\mu = \frac{\delta F[n(\mathbf{r})]}{\delta n(\mathbf{r})} + v_{\rm ext}(\mathbf{r})\,, \tag{1.43}$$

where μ is the Lagrange multiplier associated with the constraint of constant N. The idea of Kohn and Sham was to set up a system where the kinetic energy could be determined exactly, since this was a major problem in Thomas-Fermi theory. This was achieved by invoking a non-interacting system of electrons. The corresponding groundstate wavefunction $\Psi_{\rm KS}$ for this type of system is given exactly by a determinant of single-particle orbitals $\psi_i(\mathbf{r}_i)$,

$$\Psi_{\rm KS} = \frac{1}{\sqrt{N!}}\det[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\ldots\psi_N(\mathbf{r}_N)]\,. \tag{1.44}$$

The universal functional $F[n(\mathbf{r})]$ was then partitioned into three terms, the first two of which are known exactly and constitute the majority of the energy, the third being a small unknown quantity,

$$F[n(\mathbf{r})] = T_{\rm s}[n(\mathbf{r})] + E_{\rm H}[n(\mathbf{r})] + E_{\rm XC}[n(\mathbf{r})]\,. \tag{1.45}$$

$T_{\rm s}[n(\mathbf{r})]$ is the kinetic energy of a non-interacting electron gas of density $n(\mathbf{r})$, $E_{\rm H}[n(\mathbf{r})]$ is the classical electrostatic (Hartree) energy of the electrons,

$$E_{\rm H}[n(\mathbf{r})] = \frac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\,d\mathbf{r}\,d\mathbf{r}'\,, \tag{1.46}$$

and $E_{\rm XC}[n(\mathbf{r})]$ is the exchange-correlation energy, which contains the difference between the exact and non-interacting kinetic energies and also the non-classical contribution to the electron-electron interactions, of which the exchange energy is a part. In the Kohn-Sham prescription the Euler equation given in (1.43) now becomes,

$$\mu = \frac{\delta T_{\rm s}[n(\mathbf{r})]}{\delta n(\mathbf{r})} + v_{\rm KS}(\mathbf{r})\,, \tag{1.47}$$

where the Kohn-Sham potential $v_{\rm KS}(\mathbf{r})$ is given by,

$$v_{\rm KS}(\mathbf{r}) = v_{\rm ext}(\mathbf{r}) + v_{\rm H}(\mathbf{r}) + v_{\rm XC}(\mathbf{r})\,, \tag{1.48}$$

with the Hartree potential $v_{\rm H}(\mathbf{r})$,

$$v_{\rm H}(\mathbf{r}) = \frac{\delta E_{\rm H}[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \int \frac{n(\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\,d\mathbf{r}'\,, \tag{1.49}$$

and the exchange-correlation potential $v_{\rm XC}(\mathbf{r})$,

$$v_{\rm XC}(\mathbf{r}) = \frac{\delta E_{\rm XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})}\,. \tag{1.50}$$

The crucial point to understand in Kohn-Sham theory is that (1.47) is just a rearrangement of (1.43), so the density obtained when solving the alternative non-interacting Kohn-Sham system is the same as the exact groundstate density. The groundstate density is obtained in practice by solving the N one-electron Schrödinger equations,

$$\left[-\frac{1}{2}\nabla^2 + v_{\rm KS}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r})\,, \tag{1.51}$$

where ε_i are Lagrange multipliers corresponding to the orthonormality of the N single-particle states $\psi_i(\mathbf{r})$, and the density is constructed from,

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2\,. \tag{1.52}$$

The non-interacting kinetic energy $T_{\rm s}[n(\mathbf{r})]$ is therefore given by,

$$T_{\rm s}[n(\mathbf{r})] = -\frac{1}{2}\sum_{i=1}^{N}\int \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})\,d\mathbf{r}\,. \tag{1.53}$$

Since $v_{\rm KS}(\mathbf{r})$ depends on the density through the exchange-correlation potential, relations (1.48), (1.51) and (1.52), which are known as the Kohn-Sham equations, must be solved self-consistently as in the Hartree-Fock scheme described in Sec. 1.4.1.

In order to handle the kinetic energy in an exact manner, N equations have to be solved in Kohn-Sham theory to obtain the set of Lagrange multipliers $\{\varepsilon_i\}$, as opposed to one equation that determines μ when solving for the density directly, as in the Thomas-Fermi approach. However an advantage of the Kohn-Sham method is that as the complexity of a system increases, due to N increasing, the problem becomes no more difficult, only the number of single-particle equations to be solved increases.

Although exact in principle, Kohn-Sham theory is approximate in practice because of the unknown exchange-correlation functional $E_{\rm XC}[n(\mathbf{r})]$. An implicit definition of $E_{\rm XC}[n(\mathbf{r})]$ can be given through (1.45) as,

$$E_{\rm XC}[n(\mathbf{r})] = T[n(\mathbf{r})] - T_{\rm s}[n(\mathbf{r})] + E_{\rm ee}[n(\mathbf{r})] - E_{\rm H}[n(\mathbf{r})] \tag{1.54}$$

where $T[n(\mathbf{r})]$ and $E_{\rm ee}[n(\mathbf{r})]$ are the exact kinetic and electron-electron interaction energies respectively. The intention of Kohn and Sham was to make the unknown contribution to the total energy of the non-interacting system as small as possible, and this is indeed the case with the exchange-correlation energy, however it is still an important contribution since the binding energy of many systems is about the same size as $E_{\rm XC}[n(\mathbf{r})]$, so an accurate description of exchange and correlation is crucial for the prediction of binding properties. Present approximations for the exchange-correlation energy are far from satisfactory, consequently the development of improved exchange-correlation functionals is essential. An in-depth discussion of the exact properties of $E_{\rm XC}[n(\mathbf{r})]$, and the approximations presently used will be discussed in Chapter 2. Before this is presented, the remainder of this chapter concentrates on the implementation of Kohn-Sham theory for periodic systems.

1.6 Plane-Wave Implementation of DFT

This section describes the plane-wave pseudopotential implementation of Kohn-Sham DFT that is used in the calculations performed in this work. This method is well established within the physics community as it is particularly suited to describing infinite periodic systems such as solids.

1.6.1 Bloch's Theorem

Bloch's theorem [20] states that the wavefunction of an electron $\psi_{j,\mathbf{k}}$, within a periodic potential, can be written as the product of a lattice periodic part $u_j(\mathbf{r})$ and a wavelike part $\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}$,

$$\psi_{j,\mathbf{k}}(\mathbf{r}) = u_j(\mathbf{r})\,\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}\,, \tag{1.55}$$

where the subscript j indicates the band index and $\mathbf{k}$ is a continuous wavevector that is confined to the first Brillouin zone of the reciprocal lattice [20]. Since $u_j(\mathbf{r})$ has the same periodicity as the direct lattice, it can be expressed in terms of a discrete plane-wave basis set with wavevectors $\mathbf{G}$ that are reciprocal lattice vectors of the crystal, i.e.

$$u_j(\mathbf{r}) = \sum_{\mathbf{G}} c_{j,\mathbf{G}}\,\mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}\,, \tag{1.56}$$

where $\mathbf{G}\cdot\mathbf{R} = 2\pi m$, where m is an integer, $\mathbf{R}$ are the crystal lattice vectors and $c_{j,\mathbf{G}}$ are the plane-wave coefficients. The above results show that the electron wavefunctions can be expanded in terms of a linear combination of plane-waves,

$$\psi_{j,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{j,\mathbf{k}+\mathbf{G}}\,\mathrm{e}^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\,. \tag{1.57}$$

Plane-waves are a simple way of representing electron wavefunctions. They offer a complete basis set that is independent of the type of crystal and treats all areas of space equally. This is in contrast to some other basis sets which use localised functions such as Gaussians which are dependent on the positions of the ions.

1.6.2 Kohn-Sham Equations in Plane-Wave Form

Using a plane-wave basis set to expand the electronic wavefunctions in periodic systems leads to a particularly simple formulation of the Kohn-Sham equations in DFT. Accounting for the fact that the various contributions to the local potential in the Kohn-Sham equation (1.48) can be written in the form,

$$v(\mathbf{r}) = \sum_{\mathbf{G}} \bar{v}(\mathbf{G}) \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}, \tag{1.58}$$

where $\bar{v}(\mathbf{G})$ represents the Fourier transform of the corresponding real-space quantity, and substituting the plane-wave solutions given by (1.57) into (1.48), leads to a reciprocal-space representation of the Kohn-Sham equations,

$$\sum_{\mathbf{G}'} \left[\frac{1}{2}|\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{G},\mathbf{G}'} + \bar{v}_{\mathrm{ext}}(\mathbf{G}-\mathbf{G}') + \bar{v}_{\mathrm{H}}(\mathbf{G}-\mathbf{G}') + \bar{v}_{\mathrm{XC}}(\mathbf{G}-\mathbf{G}') \right] c_{j,\mathbf{k}+\mathbf{G}'} = \varepsilon_i(\mathbf{k}) c_{j,\mathbf{k}+\mathbf{G}}, \tag{1.59}$$

where it can be seen that the kinetic energy is diagonal, and the remaining three terms on the left-hand-side are the Fourier components of the external, Hartree and exchange-correlation potentials respectively. The $\mathbf{G} = 0$ component of the Hartree and external potentials diverge due to the long-range nature of the Coulomb interaction, however the divergences cancel to give a constant value that is ill-defined. However, the value of this constant can be set arbitrarily and does not affect the physical properties of a system.

For an exact calculation, the dimension of the plane-wave basis set should be infinite. Fortunately the plane-waves at the lower end of the kinetic energy range are most important, so a practical solution of (1.59) can be obtained by truncating the basis set to a finite number of plane-waves. This is defined by the kinetic cutoff energy E_{cut},

$$\frac{1}{2}|\,\mathbf{k}+\mathbf{G}\,|^2 \leq E_{\mathrm{cut}}\,. \tag{1.60}$$

These leads to another advantage of the plane-wave basis set in that the accuracy can be systematically improved by increasing E_{cut} - a feature that is certainly not shared by localised basis sets. The main disadvantage of plane-waves is that they are not efficient at describing wavefunctions with large curvature such as in the core regions of atoms, consequently such regions of space require an unreasonably large number of planes to be sufficiently accurate, and so would dominate the convergence of E_{cut}. This problem can be surmounted with the pseudopotential approximation which is described in Sec. 1.6.4.

1.6.3 k-point Sampling

By virtue of Bloch's theorem, any real-space integral over a periodic system with infinite extent can be replaced by an integral in reciprocal-space over the (finite) first Brillouin zone. However this still entails calculating the periodic functions at an infinite number of points in reciprocal space, which will be referred to as k-points. This is a consequence of the infinite number of electrons. This problem can be overcome by exploiting the fact that electron wavefunctions do not change appreciably over a small distances in k-space, therefore the integrations can be performed as summations over a finite, but sufficiently dense, mesh of k-points. So, any integrated function $f(\mathbf{r})$, such as the density or total energy, can be computed as a discrete sum,

$$\int_{\text{BZ}} F(\mathbf{k})\, d\mathbf{k} = \frac{1}{\Omega} \sum_j w_j F(\mathbf{k}_j) \tag{1.61}$$

where $F(\mathbf{k})$ is the Fourier transform of $f(\mathbf{r})$, Ω is the cell volume and w_j are weighting factors. The number of k-points required for a sufficiently accurate calculation must be ascertained by k-point sampling - a procedure in which the total energy of the system is converged with respect to increases in the k-point mesh density.

The positions of the k-points within the Brillouin zone must be carefully selected since a judicious choice will result in an efficient description of a

particular system, leading to quite significant computational savings. Different approaches for obtaining these optimal or "special" k-point sets have been discussed in the past [21, 22, 23]. However the calculations performed in this work employ the Monkhorst-Pack method [24], whereby the k-points are distributed homogeneously throughout space in rows and columns that follow the shape of the Brillouin zone, i.e.

$$\mathbf{k}_j = x_{1j}\mathbf{b}_1 + x_{2j}\mathbf{b}_2 + x_{3j}\mathbf{b}_3 \tag{1.62}$$

where $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are the reciprocal lattice vectors, and,

$$x_{ij} = \frac{l_i}{n_j}, \quad j = 1, \ldots, n_j\,. \tag{1.63}$$

where l_i are the lengths of the reciprocal lattice vectors and n_j characterises the number of special points in the set.

Typically, the point-group symmetry of the crystal is used to produce a smaller subset of the full special k-point set, containing points located within the irreducible part of the Brillouin zone. The values of the weighting factors w_j are adjusted according to this new k-point set and the integrals (1.61) are calculated with this set. This results in a significant reduction in the computational expense since a smaller number of k-points are used in the summations.

1.6.4 Pseudopotentials

Electrons in matter can be broadly categorised into two types - *core* electrons, which are strongly localised in the closed inner atomic shells, and *valence* electrons, which exist outside the core. Unfortunately, a plane-wave basis set is generally not suitable for describing electron wavefunctions since a prohibitively large number would be required to accurately describe the oscillations in the core regions which maintain orthogonality between valence and core electrons. As a result, all-electron plane-wave calculations demand a huge computational expense that is simply not practical. However, by

realising that the electronic structure of the core-electrons remains largely unchanged in different chemical environments, and is also of minimal interest generally, the problems relating to the core-electrons can be overcome by use of the pseudopotential approximation [25, 26, 27]

The pseudopotential approximation replaces the strong ionic potential $v_{\text{ion}}(\mathbf{r})$ in the core region, by a weaker pseudopotential $v_{\text{ion}}^{\text{PS}}(\mathbf{r})$. The corresponding set of pseudo-wavefunctions $\psi^{\text{PS}}(\mathbf{r})$ and the all-electron wavefunctions $\psi^{\text{AE}}(\mathbf{r})$ are identical outside a chosen cutoff radius r_c and so exhibit the same scattering properties, but $\psi^{\text{PS}}(\mathbf{r})$ does not possess the nodal structure that cause the oscillations inside r_c, which means they can now be described with a reasonable number of plane-waves. A schematic illustration of the pseudopotential concept is shown in Fig. 1.1.

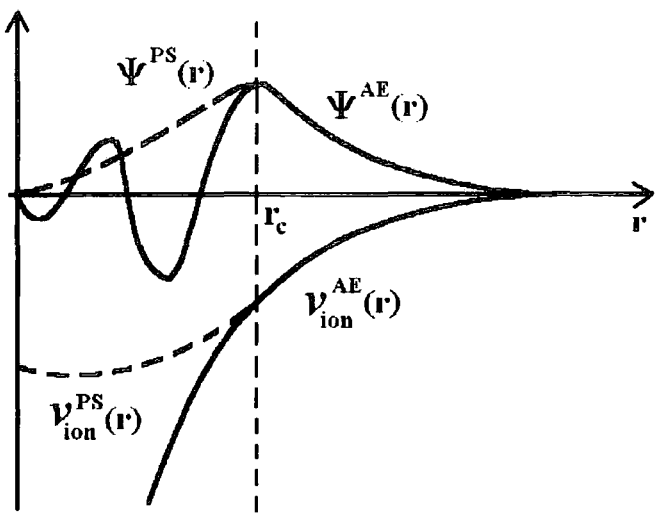

Figure 1.1: Schematic illustration of the pseudopotential concept. The solid lines show the all-electron wavefunction, $\Psi^{\text{AE}}(\mathbf{r})$, and ionic potential, $v_{\text{ion}}^{\text{AE}}(\mathbf{r})$, while the dashed lines show the corresponding pseudo-wavefunction, $\Psi^{\text{PS}}(\mathbf{r})$, given by the pseudopotential, $v_{\text{ion}}^{\text{PS}}(\mathbf{r})$. All quantities are shown as a function of distance, $\mathbf{r}$, from the atomic nucleus. The cutoff radius r_c marks the point beyond which the all-electron and pseudo quantities become identical.

1.6.4.1 First-Principles Pseudopotential Generation

The majority of pseudopotentials used in DFT calculations are generated from all-electron atomic calculations by self-consistently solving the following radial Schrödinger equation,

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - \frac{Z}{r} + v_{\rm H}(r) + v_{\rm XC}(r)\right]\psi_l^{AE}(r) = \varepsilon_l \psi_l^{AE}(r)\,, \quad (1.64)$$

where $v_{\rm H}(r)$ and $v_{\rm XC}(r)$ are the Hartree and exchange-correlation potentials, and $\psi_{n,l}^{\rm AE}$ is the all-electron atomic wavefunction with angular momentum component l. Conventionally, the pseudopotential is then constructed by satisfying four general criteria: (i) the valence pseudo-wavefunction $\psi_l^{\rm PS}(r)$ must be the same as $\psi_l^{\rm AE}(r)$ outside a given cutoff radius r_c, (ii) the charge enclosed within r_c must be equal for the two wavefunctions,

$$\int_0^{r_c} |\psi_l^{\rm PS}(r)|^2\, dr = \int_0^{r_c} |\psi_l^{\rm AE}(r)|^2\, dr\,, \quad (1.65)$$

and is normalised such that, $\int_0^\infty |\psi_l^{\rm PS}(r)|^2\, dr = \int_0^\infty |\psi_l^{\rm AE}(r)|^2\, dr = 1$. This is commonly referred to as norm-conservation. (iii) $\psi_l^{\rm PS}(r)$ must not contain any nodes and be continuous at r_c, as well as its first and second derivatives. Finally, (iv) the valence all-electron and pseudopotential eigenvalues must be equal.

The pseudopotential is not unique construction, indeed the above conditions permit a considerable amount of freedom when generating pseudo-wavefunctions, consequently many different ways have been developed for constructing pseudopotentials. Once a particular pseudo-wavefunction is created, the ionic pseudopotential is then obtained by inverting the radial Schrödinger equation (1.64), giving,

$$v_{{\rm ion},l}^{\rm PS}(r) = \varepsilon_l - v_{\rm H}^{\rm PS}(r) - v_{\rm XC}^{\rm PS}(r) - \frac{l(l+1)}{2r^2} + \frac{1}{2\psi_l^{\rm PS}(r)}\frac{d^2}{dr^2}\psi_l^{\rm PS}(r) \quad (1.66)$$

where $v_{\rm H}^{\rm PS}(r)$ and $v_{\rm XC}^{\rm PS}(r)$ are calculated from the pseudo-wavefunctions. A consequence of this procedure is that a separate pseudopotential must be generated for each angular momentum component l - i.e. the pseudopotential

is referred to as being "non-local". The pseudopotential operator $\hat{V}^{\mathrm{PS}}(r)$ can be written in a "semi-local" form [28, 29] as,

$$\hat{V}^{\mathrm{PS}}_{\mathrm{ion}}(r) = v^{\mathrm{PS}}_{\mathrm{LOC}}(r) + \sum_{l} \delta v^{\mathrm{PS}}_{l}(r)\hat{P}_{l}\,, \tag{1.67}$$

where $v^{\mathrm{PS}}_{\mathrm{LOC}}(r)$ is a local potential and $\hat{P}_l$ projects out the lth angular momentum component of the semi-local part, $\delta v^{\mathrm{PS}}_{l}(r)$,

$$\delta v^{\mathrm{PS}}_{l}(r) = v^{\mathrm{PS}}_{\mathrm{ion},l}(r) - v^{\mathrm{PS}}_{\mathrm{LOC}}(r)\,. \tag{1.68}$$

As a further generalisation of the pseudopotential procedure, Kleinman and Bylander (KB) [30] observed that greater efficiency could be attained if the non-locality was not restricted to the angular momentum part, but if the radial component was also converted into a separable non-local form. Therefore in the Kleinman-Bylander approach, the semi-local form (1.68) is converted into the fully non-local form $\delta v^{\mathrm{PS}}_{\mathrm{KB},l}(r)$, given by,

$$\delta v^{\mathrm{PS}}_{\mathrm{KB},l}(r) = \frac{|\delta v^{\mathrm{PS}}_{l}(r)\Phi^{0}_{l}(r)\rangle\,\langle\Phi^{0}_{l}(r)\delta v^{\mathrm{PS}}_{l}(r)|}{\langle\Phi^{0}_{l}(r)|\delta v^{\mathrm{PS}}_{l}(r)|\Phi^{0}_{l}(r)\rangle}\,, \tag{1.69}$$

where $\Phi^{0}_{l}(r)$ are the atomic pseudo-wavefunctions calculated with $\delta v^{\mathrm{PS}}_{l}(r)$. The Kleinman-Bylander form drastically reduces the computational resources in a pseudopotential calculation: for a plane-wave expansion of dimensionality N_{pw}, the semi-local form requires storage of $\sim (N^2_{\mathrm{pw}} + N_{\mathrm{pw}})/2$ projections for each angular momentum state, whereas the corresponding KB-pseudopotential evaluates just $\sim N_{\mathrm{pw}}$ projections and simple multiplications.

1.6.4.2 Ultrasoft Pseudopotentials

Many modern pseudopotential calculations use a generalisation of the Kleinman-Bylander form known as "ultrasoft" pseudopotentials, which were developed by Vanderbilt in the early 1990s [31]. As the name suggests, ultrasoft pseudopotentials attain much smoother (softer) pseudo-wavefunctions so use considerably fewer plane-waves for calculations of the same accuracy. This is achieved by relaxing the norm-conservation constraint (1.65), which offers

greater flexibility in the construction of the pseudo-wavefunctions. In this scheme the total valence density $n(\mathbf{r})$ is partitioned into so-called hard and soft contributions,

$$n(\mathbf{r}) = \sum_n \left[|\phi_n(\mathbf{r})|^2 + \sum_{ij} Q_{ij}(\mathbf{r}) \langle \phi_n | \beta_j \rangle \langle \beta_i | \phi_n \rangle \right] , \tag{1.70}$$

where β_i are projector functions that depend on the ionic positions, and the augmentation function $Q_{ij}(\mathbf{r})$ is given by

$$Q_{ij}(\mathbf{r}) = \psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) - \phi_i^*(\mathbf{r})\phi_j(\mathbf{r}) , \tag{1.71}$$

$\psi_i(\mathbf{r})$ are the all-electron wavefunctions, and $\phi_i(\mathbf{r})$ are ultrasoft wavefunctions constructed without satisfying the norm-conservation condition $Q_{ij}(\mathbf{r}) = 0$. Also, the orthonormality condition takes on a generalised form,

$$\langle \phi_i | S(\{\mathbf{R}_I\}) | \phi_j \rangle = \delta_{ij} , \tag{1.72}$$

where $S(\mathbf{R}_I)$ depends on the ionic positions through $|\beta_i\rangle$ and is defined as,

$$S = 1 + \sum_{ij} q_{ij} |\beta_j\rangle \langle \beta_i| , \tag{1.73}$$

with,

$$q_{ij} = \int Q_{ij}(\mathbf{r})\, d\mathbf{r} . \tag{1.74}$$

Typically, the cutoff energy E_{cut} when using ultrasofts is about half that of conventional norm-conserving pseudopotentials, for simple estimates the number of plane-waves scales as $E_{\text{cut}}^{\frac{3}{2}}$, therefore approximately one-third less plane waves are required in a given calculation. Ultrasoft pseudopotentials are used in all calculations in this work where atoms are involved.

1.6.5 Energy Minimisation

In order to calculate the physical properties of a system, i.e. for a given ionic configuration, the electronic states that minimise the Kohn-Sham total energy must always be determined. To achieve this, the general idea is

to iteratively improve upon a trial single-particle wavefunction, and whilst maintaining orthogonality with all other bands, minimise the contribution to the total energy from the current band with respect to changes in the plane-wave coefficients. Different methods can be used to perform the minimisation process.

1.6.5.1 Steepest Descents

A simple minimisation method is to move along directions of *steepest descent*, i.e. locating the minimum along the path of steepest descent and calculating new directions from the minima until the groundstate is found. The steepest descent method is however limited by the fact that each direction is chosen using information regarding the present sampling point only - disregarding the knowledge of previous search directions. This inefficiency means that in many cases a large number of iterations are needed, moreover, convergence in a finite number of steps is not guaranteed [32].

1.6.5.2 Conjugate Gradients

A more expedient minimisation method, and the one employed in all calculations in this work, is the conjugate gradients method [33]. Whereas in the steepest descents method all consecutive search directions are perpendicular, the conjugate gradients method combines the information from all previous directions in such a way as to create a subsequent search direction that is independent (conjugate) to all previous directions. In other words, the set of search directions form a linearly independent set. This achieved with the storage of only the previous search direction, rather than all previous directions as would be expected. The conjugate gradients method also guarantees that the minimum of an n-dimensional vector-space will be determined in n iterations, since each step reduces the dimensionality of the problem by 1, and after n iterations the dimensionality will be zero which leaves no directions in which to minimise and the minimum has been reached.

The conjugate direction is obtained using the previous conjugate directions and the present steepest descent direction. The condition that the single-particle wavefunctions are orthogonal for each band can be maintained by choosing steepest descent vectors that are orthogonal to all other bands. This technique reduces computational expense, firstly, since only the present band changes within an iteration (rather than modifying all the bands in order to maintain orthogonality), secondly, only the change in the density from each band after each iteration need be calculated.

1.6.5.3 Preconditioning

Typically, the number of plane-wave coefficients is of the order $\sim 10^5$ so a conventional conjugate gradient minimisation may take a long time for each band, also the efficiency of this method reduces as the kinetic energy cutoff (and hence N_{pw}) increases. To overcome these shortcomings, the conjugate gradient method is usually applied with a preconditioning scheme [34]. The Kohn-Sham Hamiltonian is considered ill-conditioned because of the broad spectrum of eigenvalues which result from the wide range of energies associated with the basis states - the plane-waves with high kinetic energy tend to dominate search directions even though the corresponding wavefunction coefficient is small. Preconditioning essentially divides the coefficients of the high kinetic energy waves by their kinetic energy, leaving the low kinetic energy waves untouched. This compresses the range of eigenvalue energies and leads to a faster rate of convergence.

The preconditioned conjugate gradient minimisation scheme achieves excellent rate of convergence in practice. Typically, only a few tens of iterations are required to converge total energies to within a satisfactory tolerance, for a plane-wave basis set containing up to $\sim 10^6$ basis functions. The calculations performed in this work use the preconditioning scheme of Teter *et al.* described in Ref. [35].

1.7 The CASTEP Code

All of the calculations reported in this thesis were performed using the plane-wave pseudopotential techniques described in Sec. 1.6. The calculations reported in Chapter 3 were carried-out with the CASTEP package developed by Payne *et al.* Ref. [32], which is written in FORTRAN 77. The calculations performed in Chapters 4 to 6 were performed using the new version of the CASTEP program, developed by Segall *et al.* and fully described in Ref. [36], which uses essentially the same techniques as the previous version but is implemented in FORTRAN 90.

Chapter 2

Exchange and Correlation

2.1 Fundamentals

This section elaborates on the known properties of exchange and correlation in DFT. The most important concept is that of the exchange-correlation hole $n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$, which is a quantum-mechanical zone surrounding every electron in an interacting system that reduces the probability of finding other electrons within the immediate vicinity. This section presents an exact definition of the exchange-correlation energy in DFT, which is given in terms of the hole, and also describes the properties of $n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$ that are used to construct approximate functionals.

2.1.1 The Exchange-Correlation Hole

The non-relativistic many-body electronic Hamiltonian for a system of N interacting electrons is given by,

$$\hat{H} = \hat{T} + \hat{V}_{\mathrm{ext}} + \hat{V}_{\mathrm{ee}} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 + \sum_{i=1}^{N} v_{\mathrm{ext}}(\mathbf{r}) + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (2.1)$$

The electron-electron interaction is a two-body operator and so the corresponding expectation value can be written as,

$$\langle \Psi | \hat{V}_{\mathrm{ee}} | \Psi \rangle = \frac{1}{2} \int\int \frac{P(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' d\mathbf{r}\,. \quad (2.2)$$

where Ψ is the normalised antisymmetric groundstate wavefunction of the system. The pair-density $P(\mathbf{r}, \mathbf{r}')$ gives the probability of simultaneously finding an electron at the point $\mathbf{r}$ within volume element $d\mathbf{r}$, and another electron at $\mathbf{r}'$ in volume element $d\mathbf{r}'$, among the other $N-2$ electrons in the system. Rigorously it is defined as,

$$P(\mathbf{r}, \mathbf{r}') = N(N-1) \int \cdots \int |\Psi(\mathbf{r}s, \mathbf{r}'s', \mathbf{r}_3 s_3, \ldots, \mathbf{r}_N s_N)|^2 \, d\mathbf{r}_3 s_3 \ldots d\mathbf{r}_N s_N \,. \tag{2.3}$$

The electron density is given by,

$$n(\mathbf{r}) = \frac{1}{N-1} \int P(\mathbf{r}, \mathbf{r}') \, d\mathbf{r}' \,, \tag{2.4}$$

since $\int n(\mathbf{r}) \, d\mathbf{r} = N$, this leads to the following condition,

$$\int \int P(\mathbf{r}, \mathbf{r}') \, d\mathbf{r}' d\mathbf{r} = N(N-1) \,. \tag{2.5}$$

In a classical description the motions of electrons are not correlated, so the probability of finding the pair of electrons at the points $\mathbf{r}$ and $\mathbf{r}'$ is simply given by a product of the density at the respective points, i.e.

$$P^{\text{class}}(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}) n(\mathbf{r}') \,, \tag{2.6}$$

substituting $P^{\text{class}}(\mathbf{r}, \mathbf{r}')$ into (2.2) yields the classical Coulomb repulsion, or Hartree energy. However this classical description violates (2.5). In reality, electrons obey Fermi statistics and so are kept apart quantum-mechanically by the Pauli-exclusion principle, and also from other non-classical Coulomb interactions. The effect of these exchange and correlation interactions is to reduce the classical value of the electron density at $\mathbf{r}$ due to the instantaneous position of the second electron located at $\mathbf{r}'$. Therefore each electron creates a depletion, or hole, of electron density around itself as a direct consequence of exchange-correlation effects. Taking account of the hole, the pair-density can be written as,

$$P(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}) n(\mathbf{r}') + n(\mathbf{r}) n_{\text{XC}}(\mathbf{r}, \mathbf{r}') \,, \tag{2.7}$$

where the quantum effects are accounted for by the exchange-correlation hole density, $n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$, surrounding each electron located at position $\mathbf{r}$.

From relation (2.5), the exchange-correlation hole satisfies an important normalisation condition known as a sum rule,

$$\int n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' = -1\,. \tag{2.8}$$

This implies that the exchange-correlation hole itself has a deficit of exactly one electron, therefore an electron and its hole constitute an entity with no net charge.

2.1.2 Exact Definition of Exchange and Correlation

An exact definition for the exchange-correlation energy $E_{\mathrm{xc}}[n(\mathbf{r})]$ in DFT can be derived using a method known as adiabatic connection, and is presented in Appendix A. The basic concept is that while keeping the density fixed, the non-interacting system is connected to the interacting system via a coupling-constant λ, which represents the strength of the electron-electron interaction: $\lambda = 0$ implies the non-interacting system and $\lambda = 1$ is the fully interacting system. The result is elegantly simple:

$$E_{\mathrm{xc}}[n(\mathbf{r})] = \frac{1}{2} \int n(\mathbf{r})\, d\mathbf{r} \int \frac{n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')}{|\,\mathbf{r} - \mathbf{r}'\,|}\, d\mathbf{r}'\,. \tag{2.9}$$

The exchange-correlation hole $n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$ is actually averaged over a coupling-constant dependent hole $n^{\lambda}_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$, given by,

$$n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}') = \int_0^1 n^{\lambda}_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')\, d\lambda\,. \tag{2.10}$$

A useful quantity to define from (2.9) is the exchange-correlation energy per particle, otherwise known as the energy density, $\varepsilon_{\mathrm{xc}}[n(\mathbf{r})]$,

$$\varepsilon_{\mathrm{xc}}[n(\mathbf{r})] = \frac{1}{2} \int \frac{n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')}{|\,\mathbf{r} - \mathbf{r}'\,|}\, d\mathbf{r}'\,. \tag{2.11}$$

Simply put, the electron many-body problem would be solved if $n_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$ were known exactly in analytic form.

The adiabatic connection method also provides other significant results. Most prominent is the fact that the difference between the interacting and non-interacting kinetic energy, $T_c[n(\mathbf{r})] = T[n(\mathbf{r})] - T_s[n(\mathbf{r})]$, is included within the definition (2.9). So it transforms an energy contribution that is kinetic in origin, into quantity that resembles a potential energy. It also provides the link between the electron density in DFT, and the many-body wavefunction Ψ, through the exchange-correlation hole in relations (2.3), (2.7) and (2.9).

2.1.3 Properties of the Exchange-Correlation Hole

The pair-density $P(\mathbf{r}, \mathbf{r}')$ is a probability, consequently the definition given in (2.7) with the value $P(\mathbf{r}, \mathbf{r}') = 0$, implies that,

$$n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}') \geq -n(\mathbf{r}') \tag{2.12}$$

and so the magnitude of the hole density can never be greater than the density at the site of the electron. The pair-density is also symmetric under the interchange of electron coordinates,

$$P(\mathbf{r}, \mathbf{r}') = P(\mathbf{r}', \mathbf{r})\,, \tag{2.13}$$

and this leads to,

$$n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}') = n_{\mathrm{XC}}(\mathbf{r}', \mathbf{r})\,\frac{n(\mathbf{r}')}{n(\mathbf{r})}\,. \tag{2.14}$$

This has consequences for the pair-correlation function described in the next subsection. The exchange-correlation hole can be conveniently separated into a summation of exchange and correlation contributions, also known as the Fermi and Coulomb holes,

$$n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}') = n_{\mathrm{X}}(\mathbf{r}, \mathbf{r}') + n_{\mathrm{C}}(\mathbf{r}, \mathbf{r}') \tag{2.15}$$

whereby the exchange (Fermi) hole is defined in terms of the λ-dependent hole as,

$$n_{\mathrm{X}}(\mathbf{r}, \mathbf{r}') = n_{\mathrm{XC},\lambda=0}(\mathbf{r}, \mathbf{r}') \tag{2.16}$$

and subsequently the correlation (Coulomb) hole is,

$$n_{\mathrm{C}}(\mathbf{r},\mathbf{r}') = n_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}') - n_{\mathrm{X}}(\mathbf{r},\mathbf{r}')\,. \tag{2.17}$$

The exchange hole can be defined exactly from the Hartree-Fock expression for the exchange energy, given by the last term in relation (1.26),

$$E_{\mathrm{X}} = \frac{1}{2}\int n(\mathbf{r})\,d\mathbf{r}\int \frac{n_{\mathrm{X}}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' \tag{2.18}$$

where the exchange hole, given in terms of spin orbitals, $\psi_j(\mathbf{r}s)$, is:

$$n_{\mathrm{X}}(\mathbf{r},\mathbf{r}') = -\frac{1}{n(\mathbf{r})}\sum_{s}\left[\sum_{j}^{N}|\psi_j^*(\mathbf{r}s)\psi_j(\mathbf{r}'s)|\right]^2 . \tag{2.19}$$

This leads to the following sum rule condition,

$$\int n_{\mathrm{X}}(\mathbf{r},\mathbf{r}')\,d\mathbf{r}' \;=\; -1\,, \tag{2.20}$$

and so the corresponding sum rule on the correlation hole must be,

$$\int n_{\mathrm{C}}(\mathbf{r},\mathbf{r}')\,d\mathbf{r}' \;=\; 0\,. \tag{2.21}$$

From relation (2.19), the exchange hole satisfies the inequality,

$$n_{\mathrm{X}}(\mathbf{r},\mathbf{r}') \leq 0\,. \tag{2.22}$$

The on-top hole is the value of the hole density when the inter-electron distance is zero, i.e. when $\mathbf{r} = \mathbf{r}'$. Considering exchange interactions only, relation (2.19) is zero for a pair of opposite-spin electrons, whereas for same spin electrons it leads to the exact on-top condition,

$$n_{\mathrm{X}}(\mathbf{r},\mathbf{r}) = -n(\mathbf{r})\,. \tag{2.23}$$

The sum rule and negativity constraints on the exchange hole represent stringent requirements, which together, determine a well defined spatial range for the exchange hole. A deep exchange hole, i.e. one with a large on-top value, will have a short spatial extent, and vice-versa. The corresponding sum rule on the correlation hole is a much weaker constraint in comparison, since $n_{\mathrm{C}}(\mathbf{r},\mathbf{r}')$ can have positive or negative values, and so a correlation hole length scale is less well defined.

2.1.4 The Pair-Correlation Function

Since the hole density $n_{\rm XC}(\mathbf{r},\mathbf{r}')$ describes the depletion of the average electron density $n(\mathbf{r})$ around an individual electron, it is useful to construct the hole in terms of a distribution function known as the pair-correlation function $g_{\rm XC}(\mathbf{r},\mathbf{r}')$, with the form:

$$n_{\rm XC}(\mathbf{r},\mathbf{r}') = n(\mathbf{r}')[g_{\rm XC}(\mathbf{r},\mathbf{r}') - 1]\,. \tag{2.24}$$

Strictly speaking, $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ is the average over the coupling constant λ,

$$g_{\rm XC}(\mathbf{r},\mathbf{r}') = \int_0^1 g^{\lambda}_{\rm XC}(\mathbf{r},\mathbf{r}')\,d\lambda\,, \tag{2.25}$$

and it is the λ dependence of $g^{\lambda}_{\rm XC}(\mathbf{r},\mathbf{r}')$ that defines $n^{\lambda}_{\rm XC}(\mathbf{r},\mathbf{r}')$ in (2.10).

From the definition (2.24), the pair-correlation function must tend to unity at large distance from an electron since exchange-correlation effects diminish as $\mathbf{r}' \to \infty$, and the pair density reverts to the classical description given by (2.6). Also, substituting $P(\mathbf{r},\mathbf{r}') = 0$ in (2.7) yields $g_{\rm XC}(\mathbf{r},\mathbf{r}') = 0$, which demonstrates that the pair-correlation function can be viewed as a probability function - for an electron located at $\mathbf{r}$, the probability of finding another electron at the point $\mathbf{r}'$, within an infinitessimal region $d\mathbf{r}'$ relative to the uncorrelated probability, is given by $g_{\rm XC}(\mathbf{r},\mathbf{r}')$.

The general shape of the pair-correlation function in cases where exchange or correlation interactions dominate are schematically shown in Fig. 2.1. As the exchange interaction becomes increasingly dominant, and therefore $n_{\rm XC}(\mathbf{r},\mathbf{r}') \sim n_{\rm X}(\mathbf{r},\mathbf{r}')$, the on-top pair-correlation function, $g_{\rm X}(\mathbf{r},\mathbf{r})$, approaches a value of 0.5. This is because for like spin electrons, $g_{\rm X}(\mathbf{r},\mathbf{r}) = 0$, from relation (2.23) and (2.24) with $\mathbf{r} = \mathbf{r}'$, while for unlike spins $g_{\rm X}(\mathbf{r},\mathbf{r}) = 1$ from (2.19), and so the average value is 0.5. The inclusion of correlation effects always reduces the value of $g_{\rm XC}(\mathbf{r},\mathbf{r})$, and so the following constraint holds for spin-unpolarised systems,

$$0 \leq g_{\rm XC}(\mathbf{r},\mathbf{r}) \leq \frac{1}{2}\,, \tag{2.26}$$

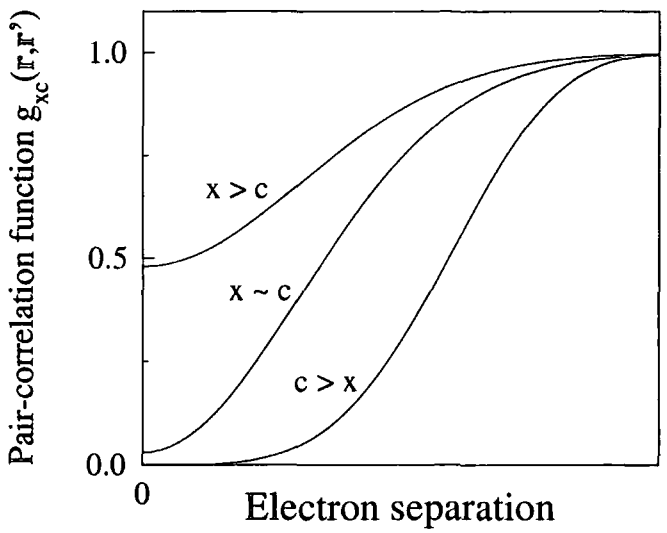

Figure 2.1: Schematic plots showing the form of the pair-correlation function $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ in situations where exchange (x) and/or correlation (c) interactions dominate.

consequently, when correlation effects are important, $g_{\rm XC}(\mathbf{r},\mathbf{r}) \to 0$, as illustrated in Fig. 2.1.

The pair-correlation function has been an important quantity of interest in many-body physics for many years. It is certainly the most fundamental element in DFT since it yields the exchange-correlation hole and therefore the total exchange-correlation energy and potential. An accurate universal model for $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ does not exist, however examples of nearly exact pair-correlation functions can be calculated for individual cases. The most frequently used methods for carrying this out are quantum Monte-Carlo techniques, which will be discussed in Sec. 2.3, since there exists a rigorous definition for $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ in terms of the coupling-constant dependent many-body wavefunction Ψ_λ; using the pair-density relations (2.3) and (2.7) together with (2.24), yields:

$$g_{\rm XC}(\mathbf{r},\mathbf{r}') = \frac{N(N-1)}{n(\mathbf{r})n(\mathbf{r}')}\int_0^1 d\lambda \times \int\cdots\int |\Psi_\lambda(\mathbf{r}s,\mathbf{r}'s',\mathbf{r}_3s_3,\ldots,\mathbf{r}_Ns_N)|^2\, d\mathbf{r}_3s_3\ldots d\mathbf{r}_Ns_N \quad (2.27)$$

Pair-correlation functions generated in this way can be used for purposes of comparison for methods that attempt to provide universally applicable models. As well as the limited number of exact conditions about $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ given previously, there are some other known constraints that can guide such models. These will be described later on this thesis.

2.1.5 Self-Interaction Effects

As well as accounting for the difference in the kinetic energy between the fully interacting system and the non-interacting Kohn-Sham scheme, exchange and correlation also cancels the self-interaction effect that originates from the Hartree term. The only case in which self-interaction can be fully defined is for a system containing exactly one electron. In this case let one of the spin densities be zero, say $n_\beta(\mathbf{r}) = 0$, then,

$$\int n_\alpha(\mathbf{r})\, d\mathbf{r} = \int n(\mathbf{r})\, d\mathbf{r} = 1\,. \tag{2.28}$$

In this instance $T_s[n(\mathbf{r})]$ and $v_{\rm ext}(\mathbf{r})$ are the exact kinetic and potential energies of the the system, and as there are no electron-electron interactions $P(\mathbf{r},\mathbf{r}') = 0$. Since the probability of finding another electron in the system is zero, the pair-correlation function must also be zero at all points in space, which leads to the result,

$$n_{\rm XC}(\mathbf{r},\mathbf{r}') = -n(\mathbf{r}')\,, \tag{2.29}$$

and therefore the exchange-correlation energy term directly cancels the Hartree energy of the system. More specifically, since condition (2.29) arises from the non-zero form of the exchange sum rule, then,

$$E_{\rm X}[n_\alpha(\mathbf{r}),0] = -E_{\rm H}[n(\mathbf{r})]\,, \qquad E_{\rm C}[n_\alpha(\mathbf{r}),0] = 0\,, \tag{2.30}$$

so the self-energy of the electron is cancelled by exchange. Similarly the exchange and correlation potentials are,

$$v_{{\rm X},\alpha}([n_\alpha(\mathbf{r})]) = -v_{\rm H}([n(\mathbf{r})]) + C_1 \tag{2.31}$$

$$v_{{\rm C},\alpha}([n_\alpha(\mathbf{r})]) = 0 + C_2\,, \tag{2.32}$$

where the constants C_1 and C_2 arise because the exchange-correlation potential can only be defined up to a constant [37]. Relations (2.31) and (2.32) show that the electron moves in the bare external potential $v_{\text{ext}}(\mathbf{r})$. For most approximate functionals the above exact relations are not satisfied, which results in a self-interaction error E^{SIE} in the computed total energies,

$$E^{\text{SIE}} = E_{\text{H}}[n(\mathbf{r})] + E_{\text{XC}}[n_\alpha(\mathbf{r}), 0]\,. \tag{2.33}$$

This equation can be used as a measure of the degree of self-interaction exhibited by a given exchange-correlation approximation. Nearly all of the conventional functionals used in DFT possess self-interaction errors i.e. $E^{\text{SIE}} \neq 0$ for a one-electron system.

Self-interaction corrected (SIC) functionals have been devised in the past such as that of Perdew and Zunger [38]. When applied to atoms the error in the total exchange-correlation energies is greatly reduced [39, 40], and the highest occupied orbitals of isolated atoms are in better agreement with experimental ionisation energies [38], which in theory should be identical if the arbitrary constant in $v_{\text{XC}}(\mathbf{r})$ is set to zero. Problems are however encountered when applying the Perdew-Zunger SIC prescription to solids since the energy functional is not invariant under a unitary transformation of the occupied orbitals, also, a localised set of basis functions must be used since the SIC energy is zero when Bloch functions are employed. Nevertheless, examples of the method have been demonstrated in solids using localised orbitals, and the most notable result is that the electronic band-gap of wide-gap insulators [41, 42] and transition-metal monoxides [43, 44] is significantly improved.

2.2 Exchange-Correlation Approximations

Functionals essentially try to model the exchange-correlation hole. This is done with varying degrees of sophistication depending on the approach taken.

However, all functionals can be written in the following general form,

$$E_{\mathrm{XC}}[n(\mathbf{r})] = \int n(\mathbf{r})\, \varepsilon_{\mathrm{XC}}(\mathbf{r})\, d\mathbf{r}\,, \tag{2.34}$$

where $\varepsilon_{\mathrm{XC}}(\mathbf{r})$ is the exchange-correlation energy per particle, or energy density for short. Functionals can be characterised by the way in which the density surrounding each electron is sampled in order to construct $\varepsilon_{\mathrm{XC}}(\mathbf{r})$. There exist five principal types of functional that have been proposed, which are now described.

2.2.1 The Local Density Approximation

The oldest, simplest and probably the most important functional is the local density approximation (LDA), which was proposed by Hohenberg and Kohn in their original DFT paper [8]. The LDA consists of locally approximating the true exchange-correlation energy of a system by the exchange-correlation energy associated with a homogeneous electron gas of the same density. The homogeneous gas is the only system for which the form of the exchange-correlation energy is known precisely. The LDA is only dependent on the local density, and the total energy is commonly written as,

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[n(\mathbf{r})] = \int n(\mathbf{r})\, \varepsilon_{\mathrm{XC}}^{\mathrm{hom}}[n(\mathbf{r})]\, d\mathbf{r}\,, \tag{2.35}$$

where $\varepsilon_{\mathrm{XC}}^{\mathrm{hom}}[n(\mathbf{r})]$ is the exchange-correlation energy density corresponding to a homogeneous electron gas of density $n(\mathbf{r})$. The energy can be decomposed into exchange and correlation contributions, $E_{\mathrm{X}}[n(\mathbf{r})]$ is just the analytic result derived by Dirac given in (1.16), while $E_{\mathrm{C}}[n(\mathbf{r})]$ is determined from an interpolation formula [45, 46] that connects the known limiting form of $\varepsilon_{\mathrm{C}}^{\mathrm{hom}}[n(\mathbf{r})]$ in the high [47, 48] and low density limits [49]. A commonly used correlation formula is that of Perdew and Zunger [50] which uses accurate quantum Monte Carlo data of the homogeneous electron gas generated by Ceperley and Alder [51], to fix the coefficients in the interpolation formula.

Despite its simplicity, the LDA works well for solid systems and has been used in solid state calculations for many years, however its success does not transfer to chemistry. The LDA has a notorious tendency to overbind, and is particularly severe for molecules.

2.2.2 The Generalised Gradient Approximation

Hohenberg and Kohn presumed that the LDA would be too simplistic to work for real systems and so proposed an extension to the LDA known as the gradient expansion approximation (GEA) [8]. The GEA is a series expansion of increasingly higher order density gradient terms. The first order form of the GEA was subsequently implemented and tested for atoms and molecules and was a complete failure. The source of the GEA problems was later found to be caused by the violation of the sum rule (2.20) and the non-positivity constraint (2.22) on the exchange hole - both of which are important physical conditions that happen to be fulfilled by the LDA. Despite the disappointing results, the GEA provided the basis for the generalised gradient approximation (GGA) which is currently the most popular exchange-correlation functional in condensed matter physics.

The vital steps that lead to the GGA were principally made by Perdew and co-workers [52] who devised a cutoff procedure that sharply terminates the GEA exchange-correlation hole in real-space using delta functions, in order to restore the sum rule and non-positivity hole conditions. As a result of this procedure the GGA can be conveniently written in terms of an analytic function known as the enhancement factor, $F_{\rm XC}[n(\mathbf{r}), \nabla n(\mathbf{r})]$, that directly modifies the LDA energy density,

$$E_{\rm XC}^{\rm GGA}[n(\mathbf{r})] = \int n(\mathbf{r})\, \varepsilon_{\rm XC}^{\rm hom}[n(\mathbf{r})]\, F_{\rm XC}[n(\mathbf{r}), \nabla n(\mathbf{r})]\, d\mathbf{r}\,. \tag{2.36}$$

Usually the GGA enhancement factor is written in terms of the Seitz radius

r_s, and the dimensionless reduced density gradient $s(\mathbf{r})$,

$$s(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|}{2k_{\mathrm{F}}(\mathbf{r})n(\mathbf{r})}, \tag{2.37}$$

where k_{F} is the Fermi-wavevector,

$$k_{\mathrm{F}}(\mathbf{r}) = [3\pi^2 n(\mathbf{r})]^{1/3}. \tag{2.38}$$

Since the GGA is not a unique functional form, plotting $F_{\mathrm{XC}}(r_s, s)$ against s for various r_s values allows an effective way of examining and comparing different GGAs.

Despite the crudeness of the real-space cutoff procedure, the GGA successfully resolved the two main failures of the GEA that it was intended to correct, and also gave improvement over the LDA in several instances. The most notable outcome was the significant reduction in the LDA overbinding error for solids and molecules. It would not be unfair to say that the success of the GGA for molecular properties was a major factor in the part awarding of the Nobel Prize in chemistry to Kohn in 1998.

An important GGA functional used predominantly by the solid state DFT community, and also in Chapter 3 of this work, is that of Perdew and Wang, known as PW91 [53, 54]. PW91 is an example of a non-empirical construction since it does not contain any free parameters that are fitted to experimental data, rather it is determined from exact quantum-mechanical relations. The exchange enhancement factor has the form,

$$F_{\mathrm{X}}^{\mathrm{PW91}}(s) = \frac{1 + 0.19645s\,\sinh^{-1}(7.7956s) + (0.2743 - 0.15084\,\mathrm{e}^{-100s^2})s^2}{1 + 0.19645s\,\sinh^{-1}(7.7956s) + 0.004s^4}, \tag{2.39}$$

which is an extension of a form given by Becke known as B88 [55], although it is tailored in order to obey extra exact conditions such as the correct behaviour in the slowly varying (small s) limit, some scaling relations [56], and energy bounds [57]. It should be noted that there is no r_s dependence in the GGA exchange enhancement factor since the exchange energy scales

linearly with uniform density scaling [56] and therefore the form of $F_{\rm X}(s)$ remains unchanged with different r_s values.

The mathematics of GGA correlation terms are complicated by the scaling relations and the different interactions that occur between like and unlike spin components. The spin-compensated PW91 correlation energy can be written as,

$$E_{\rm C}^{\rm PW91}[n_\alpha, n_\beta] = \int n(\mathbf{r})[\varepsilon_{\rm C}(r_s,\zeta) + H(t, r_s, \zeta)]\, d\mathbf{r}\,, \tag{2.40}$$

where $\varepsilon_{\rm C}(r_s,\zeta)$ is the Perdew-Wang parametrisation of the homogeneous electron gas correlation energy [58], and t is another dimensionless gradient term given by,

$$t = \frac{|\nabla n(\mathbf{r})|}{2gk_s n(\mathbf{r})}\,, \tag{2.41}$$

where $k_s = (4k_{\rm F}/\pi)^{1/2}$, $g = [(1+\zeta)^{2/3} + (1-\zeta)^{2/3}]/2$ and ζ is the degree of spin-polarisation. The function $H = H_0 + H_1$, is defined as,

$$H_0 = g^3\frac{\beta^2}{2\alpha}\log\Big[1 + \frac{2\alpha}{\beta}\frac{t^2 + At^4}{1 + At^2 + A^2t^4}\Big]\,, \tag{2.42}$$

$$H_1 = \nu\Big[C_c(r_s) - C_c(0) - \frac{3}{7}C_x\Big]g^3t^2\, e^{[-100g^4(k_s^2/k_{\rm F}^2)t^2]}\,, \tag{2.43}$$

$$A = (2\alpha/\beta)(1/e^{-2\alpha\varepsilon/(g^3\beta^2)} - 1)\,, \tag{2.44}$$

with $\alpha = 0.09$, $\beta = \nu C(0)$, $\nu = (16/\pi)(3\pi^2)^{1/3} = 0.004235$, $C_x = -0.001667$. $C_c(r_s)$ is given by Rasolt and Geldart [59],

$$C_c(r_s) = C_{xc}(r_s) - C_x \tag{2.45}$$

$$C_{xc}(r_s) = 10^{-3}\frac{2.568 + ar_s + br_s^2}{1 + cr_s + dr_s^2 + 10br_s^3}\,, \tag{2.46}$$

with $a = 23.266, b = 7.389\times10^{-3}, c = 8.723, d = 0.472$. The rationale behind the form of $H(t, r_s, \zeta)$ is described in Ref. [53, 60]. Briefly, it is governed by the behaviour of the correlation hole in the limit of high-density [61, 62].

Until recently, PW91 was the sole GGA used by the physics community. It has now probably been superseded by a modified form devised by Perdew,

Burke and Ernzerhof known as PBE [63, 64], which uses a much simplified exchange enhancement factor of the form:

$$F_{\mathrm{X}}^{\mathrm{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa}\,, \tag{2.47}$$

where $\mu = 0.21951$ and $\kappa = 0.804$. PBE was designed to give a simpler functional form by retaining only the most energetically important conditions satisfied by PW91. However the PBE and PW91 enhancement factors are virtually indistinguishable for $0 < s < 3$, i.e. the range exhibited by most physical systems, consequently they yield essentially the same physical properties.

2.2.3 Meta-Generalised Gradient Approximation

An active line of research into functionals that go beyond the GGA at the moment is the meta-GGA (MGGA) form [65, 66]. MGGAs include additional semi-local information beyond the first-order density gradient contained in the GGA, such as higher order density gradients, or more popular is the inclusion of the kinetic energy density $\tau(\mathbf{r})$ which involves derivatives of the occupied Kohn-Sham orbitals,

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{i}^{occ} |\nabla\psi_i(\mathbf{r})|^2\,. \tag{2.48}$$

The integrated $\tau(\mathbf{r})$ is equivalent to the usual non-interacting kinetic energy $T_s[n(\mathbf{r})]$, given by relation (1.53), i.e. $T_s[n(\mathbf{r})] = \int \tau(\mathbf{r})\, d\mathbf{r}$. The MGGA may be written with the general form,

$$E_{\mathrm{XC}}^{\mathrm{MGGA}}[n(\mathbf{r})] = \int f[n(\mathbf{r}), \nabla n(\mathbf{r}), \nabla^2 n(\mathbf{r}), \tau(\mathbf{r}), \mu(\mathbf{r}), \ldots \gamma(\mathbf{r})]\, d\mathbf{r}\,, \tag{2.49}$$

where $\mu(\mathbf{r}), \ldots \gamma(\mathbf{r})$ are other possible semi-local quantities (i.e. defined locally at $\mathbf{r}$) that could be used in the construction of MGGAs.

There are several MGGA forms now in existence [67, 68, 69, 70, 71, 72, 73, 74, 75, 76] and some improvement has been obtained over the GGA in a limited number of tests [77]. However a few cautionary words should be said

about the MGGA. At present, MGGA calculations for solids are performed inconsistently because they resort to using GGA orbitals and densities to evaluate $E_{\rm XC}^{\rm MGGA}[n(\mathbf{r})]$, since the orbital dependence does not permit an easy evaluation of a multiplicative exchange-correlation potential $v_{\rm XC}(\mathbf{r})$. Therefore properties are only calculated at experimental structures. To achieve self-consistency using a multiplicative potential, computationally expensive methods such as the optimised effective potential (OEP) [78, 79] must be invoked, however this has yet to be implemented - indeed it may eventually prove too costly for practical computations. Another point to highlight is that all MGGA forms are constructed using experimental molecular data to define the form. This will have the effect of introducing an element of bias into the character of the functional. This issue is investigated in Chapter 3 with regard to GGA functionals.

2.2.4 Hybrid Functionals

An interesting class of functionals are hybrids [80], which combine exact (Hartree-Fock) exchange with conventional GGAs, the general form is,

$$E_{\rm XC}^{\rm hybrid} = \alpha(E_{\rm X}^{\rm HF} - E_{\rm X}^{\rm GGA}) + E_{\rm XC}^{\rm GGA}\,, \tag{2.50}$$

where $E_{\rm X}^{\rm HF}$ is the Hartree-Fock exchange expression given in (1.26), except Kohn-Sham rather than Hartree-Fock orbitals are used, hence the wording "exact-exchange". The coefficient, α, that determines the amount of exact-exchange mixing cannot be assigned from first-principles and so is fitted semi-empirically.

The logic behind this prescription was put forward by Becke [80] who noted that the limits of the adiabatic connection integral for the exact exchange-correlation energy (A.13) could be approximated as:

$$E_{\rm XC} = \int_0^1 U^\lambda \, d\lambda = \frac{1}{2}U^0 + \frac{1}{2}U^1\,. \tag{2.51}$$

Since $\lambda = 0$ corresponds to the exchange only limit, this could well be described using Hartree-Fock theory, while $\lambda = 1$ represents the most local

part of the electron interactions, as a result of correlation, and so could be amenable to a local-type density functional treatment. As a result, Becke proposed the so-called half-and-half functional,

$$E_{\rm XC} = \frac{1}{2}E_{\rm X}^{\rm HF} + \frac{1}{2}E_{\rm XC,\lambda=1}^{\rm DF}\,, \tag{2.52}$$

where $E_{\rm XC,\lambda=1}^{\rm DF}$ is obtained from a density functional approximation such as the LDA. It later emerged from semi-empirical fits to atomic and molecular data that the optimum amount of exchange mixing should be reduced to $\sim$ 0.25, although the precise value to employ depends upon the fitting data [81].

Hybrids give significant improvement over GGAs for many molecular properties, consequently they are a very popular choice of functional in quantum chemistry. Possibly the most widely used hybrid is the B3LYP functional proposed by Stevens *et al.* [82] which is a generalisation of the B3P86 form devised by Becke [83]. Hybrids are not generally used in solid state physics because of the difficulty of computing the exact-exchange part within a plane-wave basis set. Nonetheless, hybrid functionals successfully demonstrate the need to incorporate fully non-local information in order to deliver greater accuracy.

2.2.5 Non-Local Functionals

The final class of functionals to consider are fully non-local approximations such as the average density approximation (ADA) [84] and weighted density approximation (WDA) [85, 86, 87], which were created in the 1970s. The philosophy behind these functionals is to use the exact density functional expression for $E_{\rm XC}[n(\mathbf{r})]$ given by (2.9), and directly model the exchange-correlation hole using analytic functions. These fully non-local functionals therefore have the following general form,

$$E_{\rm XC}^{\rm NL}[n(\mathbf{r})] = \frac{1}{2}\int n(\mathbf{r})\,d\mathbf{r}\int \frac{n_{\rm XC}^{\rm model}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\,d\mathbf{r}'\,. \tag{2.53}$$

The advantage of the non-local approach is that the number and severity of the approximations are kept to a minimum, and so the functionals retain many of the correct features of the exact functional such as self-interaction effects and correct asymptotic characteristics. However, the main disadvantage is the increase in computational expense due to the double integral form of (2.53). An unfortunate consequence of this downside is that fully non-local functionals such as the ADA and WDA are relatively unknown and little explored in comparison to the functionals described previously, despite possessing several desirable features.

The WDA, more so than the ADA, has many promising characteristics that seem to outweigh the computational effort of the method, especially since conventional semi-local functionals contain deep seated problems. It is for these reasons, and also because of the greater computational power available nowadays, that the WDA is investigated for the most part in this thesis.

2.3 Quantum Monte Carlo

Quantum Monte-Carlo (QMC) methods are now mentioned briefly because of the accurate exchange-correlation data they can generate for use in functional development, such as total energies, energy densities and holes. An extensive and recent review of QMC methods is given by Foulkes *et al.* in Ref. [88].

Whereas Hartree-Fock theory and DFT are mean-field theories in that they invoke a single-particle description which replaces the real forces of interaction between electrons with an averaged or mean-field, QMC takes the alternative approach and computes the actual many-electron wavefunction Ψ for the system. There are two principle QMC methods - variational Monte-Carlo (VMC) and diffusion Monte-Carlo (DMC). VMC is the cheapest and less accurate of the two methods, but exchange-correlation data is more readily available from this method. A VMC simulation proceeds by first choosing

a trial wavefunction Ψ_T, which contains functions representing the electron-nuclear and electron-electron correlations. These functions are then adjusted so as to minimise the variance of the total energy according to the variational principle. So the accuracy of the VMC method depends on how well the form of Ψ_T can represent the particular system.

The DMC method [89, 90, 91] is the most accurate groundstate electronic structure method, at least for extended systems. The only approximation in DMC is the location of the nodes of the wavefunction, i.e. where Ψ equals zero and changes sign. These are fixed throughout a simulation and the wavefunction is optimised between the nodes. This is commonly called the fixed-node approximation [90]. Usually the nodes from VMC wavefunctions are used as the input. The DMC method involves solving the imaginary time many-electron Schrödinger equation using a population of "walkers" that randomly sample the $3N$-dimensional vector space - the groundstate wavefunction can be obtained from the population density of the walkers after a sufficient amount of imaginary time has elapsed.

Probably the most important DMC simulations were performed for the homogeneous electron gas by Ceperley and Alder in 1980 [51], since this lead to the accurate determination of the parameters in the correlation part of the LDA. As a consequence, QMC essentially made DFT practical. However, a symbiotic relationship exists between DFT and the QMC methods nowadays: although the DMC method effectively brought about the LDA - which forms the basis of all functionals - the majority of QMC calculations currently performed, use densities and even pseudopotentials generated from DFT - usually with the LDA.

2.4 Summary

The five approximations discussed: LDA, GGA, MGGA, hybrid and WDA, represent the main functionals that have been proposed. Since the aim of

Table 2.1: Summary of the main characteristics of the different exchange-correlation functionals and which exact properties are obeyed. The last column refers to the exact functional defined by (2.9).

Property	LDA	GGA	MGGA	Hybrid	WDA	Exact
Non-empirical	Yes	Yes†	No	No	Yes	–
Locality††	L	SL	SL	NL/SL	NL	NL
Explicit local XC-hole	Yes	No	No	No	Yes	Yes
Explicit energy density $\varepsilon_{\rm XC}(\mathbf{r})$	Yes	No	No	No	Yes	Yes
$\lim_{r\to\infty}\varepsilon_{\rm XC}(\mathbf{r})$	$-e^{-\alpha r}$	$-e^{-\alpha r}$	$-e^{-\alpha r}$	$-\frac{1}{2r}$	$-\frac{1}{2r}$	$-\frac{1}{2r}$
$\lim_{r\to\infty} v_{\rm XC}(\mathbf{r})$	$-e^{-\beta r}$	$-e^{-\beta r}$	$-e^{-\beta r}$	N/A	$-\frac{1}{2r}$	$-\frac{1}{r}$
HEG limit	Yes	Yes	Yes	No	Yes	Yes
Self-interaction correction	No	No	No	No	Yes‡	Yes

† Note GGAs can also be defined semi-empirically
†† L = local, SL = semi-local, NL = non-local
‡This is true only in principle.

this thesis is to improve the current status of functional development, it is first necessary to ascertain what the best type of functional is to develop. To help make this decision, a summary of the five types is given in Table 2.1, which compares the main features of each functional, with specific reference to principle exact conditions. The last column entitled "Exact" refers to the exact functional (2.9) given by the adiabatic connection method.

It is clear from this table that the WDA seems the logical choice for functional development since it obeys the most exact conditions. Another factor is that a fully self-consistent version of the WDA is implementable within a plane-wave code, which is difficult for the MGGA and hybrid functionals.

As a consequence, the WDA will form a major part of this thesis. However before the WDA is investigated, a highly flexible GGA containing 15 fitted parameters is studies in the next chapter to assess whether the GGA form has actually reached the limits of its accuracy.

Chapter 3

Assessment of Semi-Empirical GGA Functionals

3.1 Introduction

Two distinct philosophies have emerged in the construction of modern exchange-correlation functionals. Perdew [92, 93] supports the idea that functionals should be derived non-empirically using rigorous quantum-mechanical principles and exact conditions, however Becke [94] advocates the semi-empirical approach whereby a general functional form containing free parameters is proposed, and the parameters are subsequently fitted to minimise the error in exact physical properties. The semi-empirical concept is extensively used and developed within the quantum chemistry community where there is a wealth of known atomic and molecular data [95, 96] that can be used to fit functionals.

Non-empirical GGAs such as PW91 and PBE are typically more accurate for solid state properties than their semi-empirical counterparts, and so are used more widely by the physics community [97]. However, it is important to ascertain whether recent developments in semi-empirical molecular GGAs, which have seldom been applied to extended systems, can provide an

improvement for solid state predictions. A recently proposed functional that attempts to test the limits of the semi-empirical GGA concept is the HCTH form devised by Hamprecht, Cohen, Tozer, and Handy [98]. This functional is highly flexible since it contains fifteen free parameters, and has proved successful in molecular studies [99, 100, 101]. Consequently HCTH is a good candidate to test whether semi-empirical GGAs can be used successfully in solid state applications.

3.2 The HCTH Functional

A standardised set of experimental data, collated by Pople and co-workers and known as the Gaussian 2 (G2) set [102], was especially suited for the purposes of constructing semi-empirical functionals. The G2 set consists of highly accurate experimental thermochemical data - atomisation energies, ionisation potentials and electron and proton affinities - of a range of atomic and molecular systems, drawn from the first two rows of the periodic table. Whereas some semi-empirical functionals are fitted just to this data [95, 96], the HCTH class of functional is different in that the fit data consists of the G2 set and also accurate exchange-correlation potentials calculated from high-level *ab initio* densities [103]. The advantage of this extra data is that it includes local exchange-correlation information, i.e. on a point-wise basis, as opposed to only integrated quantities such as total energies or energy differences that are contained in the G2 set.

Care must be taken to ensure that semi-empirical functionals do not become "over-fitted" do the test data. This fact was taken into account in the construction of HCTH - the use of extra exchange-correlation potential data in HCTH therefore increased the optimum number of parameters that can be otherwise employed.

3.2.1 The General Form

The HCTH functional form is an expansion of the GGA part of the B97 hybrid-exchange functional [81], with a general spin-polarised form:

$$E_{\mathrm{XC}}^{\mathrm{HCTH}}[n_\sigma(\mathbf{r})] = \int f_{\mathrm{XC}}[n_\sigma(\mathbf{r}), |\nabla n_\sigma(\mathbf{r})|]\, d\mathbf{r}\,. \tag{3.1}$$

where $\sigma = \alpha, \beta$ represents the spins of the electrons. As with all GGA functionals, the total energy expression is split into exchange and correlation contributions,

$$E_{\mathrm{XC}}^{\mathrm{HCTH}}[n_\sigma(\mathbf{r})] = E_{\mathrm{X}}[n_\sigma(\mathbf{r})] + E_{\mathrm{C}}[n_\sigma(\mathbf{r})]\,. \tag{3.2}$$

The exchange part is given by,

$$E_{\mathrm{X}}[n_\sigma(\mathbf{r})] = \sum_\sigma \int e_{\mathrm{X}\sigma}^{\mathrm{LSDA}}[n_\sigma(\mathbf{r})]\, g_{\mathrm{X}\sigma}(s_\sigma)\, d\mathbf{r} \tag{3.3}$$

$$g_{\mathrm{X}\sigma}(s_\sigma) = \sum_{i=0}^{m} c_{\mathrm{X}\sigma,i}\, u_{\mathrm{X}\sigma}^i(s_\sigma) \tag{3.4}$$

$$u_{\mathrm{X}\sigma}^i(s_\sigma) = \gamma_{\mathrm{X}\sigma} s_\sigma^2 (1 + \gamma_{\mathrm{X}\sigma} s_\sigma^2)^{-1} \tag{3.5}$$

$$\gamma_{\mathrm{X}\sigma} = 0.004\,, \tag{3.6}$$

where m is the maximum number of terms in the expansion, and,

$$s_\sigma^2 = |\nabla n_\sigma(\mathbf{r})|^2 n_\sigma^{-8/3}(\mathbf{r})\,. \tag{3.7}$$

The correlation term is composed of a parallel-spin $E_{\mathrm{C}\sigma\sigma}$ and anti-parallel spin $E_{\mathrm{C}\alpha\beta}$ components:

$$E_{\mathrm{C}}[n_\sigma(\mathbf{r})] = \sum_\sigma E_{\mathrm{C}\sigma\sigma}[n_\sigma(\mathbf{r})] + E_{\mathrm{C}\alpha\beta}[n_\sigma(\mathbf{r})]\,, \tag{3.8}$$

where the parallel term has the form,

$$E_{\mathrm{C}\sigma\sigma}[n_\sigma(\mathbf{r})] = \int e_{\mathrm{C}\sigma\sigma}^{\mathrm{LSDA}}[n_\sigma(\mathbf{r})]\, g_{\mathrm{C}\sigma\sigma}(s_\sigma)\, d\mathbf{r} \tag{3.9}$$

$$g_{\mathrm{C}\sigma\sigma}(s_\sigma) = \sum_{i=0}^{m} C_{\mathrm{C}\sigma\sigma,i}\, u_{\mathrm{C}\sigma\sigma}^i(s_\sigma) \tag{3.10}$$

$$u_{\mathrm{C}\sigma\sigma}^i(s_\sigma) = \gamma_{\mathrm{C}\sigma\sigma} s_\sigma^2 (1 + \gamma_{\mathrm{C}\sigma\sigma} s_\sigma^2)^{-1} \tag{3.11}$$

$$\gamma_{\mathrm{C}\sigma} = 0.2\,, \tag{3.12}$$

and anti-parallel spin correlations are described by,

$$E_{C\alpha\beta}[n_\sigma(\mathbf{r})] = \int e^{\mathrm{LSDA}}_{C\alpha\beta}[n_\sigma(\mathbf{r})]\, g_{C\alpha\beta}(s_{avg})\, d\mathbf{r} \tag{3.13}$$

$$g_{C\alpha\beta}(s_{avg}) = \sum_{i=0}^{m} C_{C\alpha\beta,i}\, u^i_{C\alpha\beta}(s_{avg}) \tag{3.14}$$

$$u^i_{C\alpha\beta}(s_{avg}) = \gamma_{C\alpha\beta} s^2_{avg}(1 + \gamma_{C\alpha\beta} s^2_{avg})^{-1} \tag{3.15}$$

$$\gamma_{C\alpha\beta} = 0.006\,, \tag{3.16}$$

with $s^2_{avg} = (s^2_\alpha + s^2_\beta)/2$. The quantities $\varepsilon^{\mathrm{LSDA}}_{C\sigma\sigma}$, $\varepsilon^{\mathrm{LSDA}}_{C\alpha\beta}$ are obtained from the homogeneous electron gas parametrisation of Perdew and Wang [58], using the procedure defined by Stoll *et al.* [104].

So there are three coefficients associated with each term, i, in the series expansion - one for the exchange $C_{X\sigma,i}$ contribution, and two for correlation $C_{C\sigma\sigma,i}$, $C_{C\alpha\beta,i}$. The optimal value $M = 4$ was determined from the fit data, leading to a total of 15 free parameters. The coefficients defining HCTH are given in Appendix B.

3.2.2 Details of the Implementation

The HCTH functional was implemented within the CASTEP code [32]. For full self-consistency, the exchange-correlation energy and potential must both be constructed explicitly. The spin-dependent exchange-correlation potential $v_{XC,\sigma}(\mathbf{r})$, associated with definition (3.1) is given by,

$$\begin{aligned} v_{XC,\sigma}(\mathbf{r}) &= \frac{\partial f_{XC}}{\partial n_\sigma(\mathbf{r})} - \nabla . \frac{\partial f_{XC}}{\partial \nabla n_\sigma(\mathbf{r})} \\ &= \frac{\partial f_{XC}}{\partial n_\sigma(\mathbf{r})} - \nabla . \left(\frac{\partial f_{XC}}{\partial |\nabla n_\sigma(\mathbf{r})|} \frac{\partial |\nabla n_\sigma(\mathbf{r})|}{\partial \nabla n_\sigma(\mathbf{r})} \right), \end{aligned} \tag{3.17}$$

where the first term is the LDA result, and the second originates from the use of density gradients. Since,

$$\frac{\partial |\nabla n_\sigma(\mathbf{r})|}{\partial \nabla n_\sigma(\mathbf{r})} = \frac{\nabla n_\sigma(\mathbf{r})}{|\nabla n_\sigma(\mathbf{r})|} \tag{3.18}$$

the potential (3.17) can be expressed as,

$$v_{XC,\sigma}(\mathbf{r}) = \frac{\partial f_{XC}}{\partial n_\sigma(\mathbf{r})} - \nabla . \left(\frac{\partial f_{XC}}{\partial |\nabla n_\sigma(\mathbf{r})|} \frac{\nabla n_\sigma(\mathbf{r})}{|\nabla n_\sigma(\mathbf{r})|} \right), \tag{3.19}$$

using shorthand notation for f_{XC}. The use of plane waves in the CASTEP code makes the evaluation of a GGA potential straightforward. The derivative of the density can be trivially evaluated as,

$$\nabla n(\mathbf{r}) = \sum_{\mathbf{G}} i\,\mathbf{G}\,\bar{n}(\mathbf{G})\,\mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}, \tag{3.20}$$

where $\bar{n}(\mathbf{G})$ is the density in reciprocal-space. The second term in the potential (3.19) is then obtained in a similar fashion by Fourier transforming the bracketed term to reciprocal-space, multiplying by the appropriate reciprocal-space vector, and transforming back to real space.

3.2.3 Testing: The Hellmann-Feynman Theorem

The implementation of any self-consistent exchange-correlation functional can be checked by comparing numerical and analytic forces arising from microscopic displacements of nuclei within an arbitrary system. The agreement between the two forces demonstrates that the XC energy functional is consistent with the potential.

The analytic forces are obtained from a theorem due to Hellmann [105] and Feynman [106], which states that when the valence electron wavefunctions are variationally optimised, the physical force on an ion is simply the classical electrostatic force due to the electrons and nuclei. Due to the extensive basis sets used in plane-wave calculations, the energy is essentially variational and so the force should be given by the same Hellmann-Feynman expression. In other basis set calculations, such as those using localised Gaussian functions, the basis size is smaller and so the Hellmann-Feynman theorem does not hold. An additional term known as the Pulay force [107] must be calculated in these cases.

The HCTH implementation was checked by calculating forces within the H_2 molecule, using a bond length of 0.800Å. The numerical force $\mathbf{F}^{\mathrm{num}}$ in the molecule was determined from the displacement of one of the ions through a

distance $\delta\mathbf{r}$, from its equilibrium position $\mathbf{R}_0$,

$$\mathbf{F}^{\text{num}} = -\frac{E(\mathbf{R}_0 + \delta\mathbf{r}) - E(\mathbf{R}_0 - \delta\mathbf{r})}{2\delta\mathbf{r}}. \tag{3.21}$$

Since the Hellmann-Feynman theorem requires the wavefunctions to be variationally optimised, the best agreement between the numerical and Hellmann-Feynman forces is dictated by the completeness of the basis set. Therefore the kinetic energy cutoff, E_{cut}, of the plane-wave expansion must be sufficiently large in order to make accurate comparisons. The test system consisted of placing an H_2 molecule in a supercell of size $4 \times 2 \times 2$Å, with the bond axis along the larger cell dimension. The supercell approach allows finite systems such as isolated atoms and molecules to be treated within a periodic representation, by placing the species in the centre of a periodic cell that has a sufficient amount of vacant space in order to minimise the interactions from the equivalent species in neighbouring cells. Ordinarily, a total energy convergence test is performed with respect to the size of the supercell to ensure that these neighbouring interactions are diminished, however this is not necessary for the purpose of these tests.

Using a well converged cutoff energy, the numerical and analytic Hellmann-Feynman forces agreed to within 0.002%, for a displacement of $\delta\mathbf{r} = 0.01$Å. The fact that this is a very small discrepancy and that the same degree of accuracy was also obtained with the previously established PW91 functional in CASTEP, demonstrates the correct implementation of the HCTH functional. The spin-polarised version of HCTH was also checked using the open shell H_2^- molecule.

3.3 Properties of group IV and III-V Semiconductors

Many physical properties can be obtained from the calculation of the total energy E_{Tot} of a system, which can be used to test functionals when compared

with known experimental values. For solids, structural properties such as the lattice constant a_0 and bulk modulus B_0 are usually determined, and the cohesive energy is used to assess the energetic predictions of the functional. These quantities are therefore used to test the HCTH functional.

Results of calculations performed using the HCTH functional are now presented for several group IV and III-V insulators and semiconductors. It is noted that Kurth *et al.* [108] have already applied HCTH to determine equilibrium unit cell volumes and bulk moduli for a range of solids using the linearised augmented plane-wave (LAPW) method, however, their calculations were not self-consistent as they used densities obtained from the PBE GGA functional [63].

Ultrasoft pseudopotentials [31] have been generated for each system using HCTH, and also PW91. So all of the results presented are fully consistent. Figs 3.1(a) and 3.1(b) show the kinetic energy cutoff and k-point sampling convergence tests for Si using HCTH. This convergence is typical for the diamond and zinc-blende systems examined here. Consequently all calculations performed in this chapter use a converged kinetic energy cutoff of 300 eV, and a $4 \times 4 \times 4$ Monkhorst-Pack special k-point set for the Brillouin-zone integrations, which corresponds to 28 k-points in the irreducible wedge. The experimental results are taken from Ref. [109], unless otherwise stated.

3.3.1 Lattice Constants

The lattice constant of a solid, a_0, corresponds to the size of the conventional unit cell length at the equilibrium volume, and is obtained computationally by minimising the total energy as a function of cell volume. Experimental lattice constants are usually obtained from low temperature X-ray diffraction measurements, and extrapolated to zero Kelvin.

The calculation of theoretical lattice constants is straightforward for cubic systems - single-point energy calculations are performed at several different

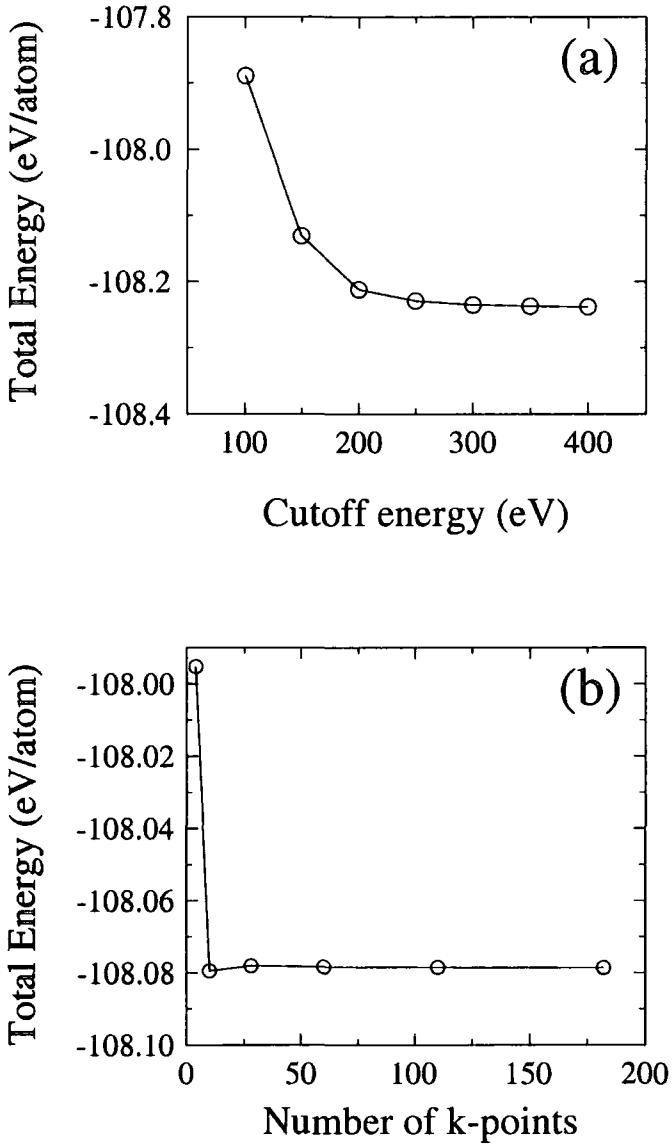

Figure 3.1: Example convergence tests are shown for bulk Si, determined using the HCTH functional. The total energy per atom (in eV) is converged with respect to (a) the kinetic energy cutoff, and (b) the number of k-points in the irreducible wedge.

volumes using the same kinetic energy cutoff and k-point sampling, and the results are fitted to the Murnaghan equation of state [110],

$$E = \frac{B_0 V}{B_0'(B_0'-1)}\left[B_0'\left(1-\frac{V_0}{V}\right)+\left(\frac{V_0}{V}\right)^{B_0'}-1\right]+E(V_0)\,. \qquad (3.22)$$

Here B_0 is the equilibrium bulk modulus which is defined in Sec. 3.3.2, and B_0' is the derivative of the bulk modulus. Usually 6 to 8 calculated energy-volume points that span the equilibrium volume by approximately ±5% are sufficient for an accurate determination of a_0. An example of a Murnaghan interpolation is shown in Fig. 3.2 for Si obtained using the HCTH functional.

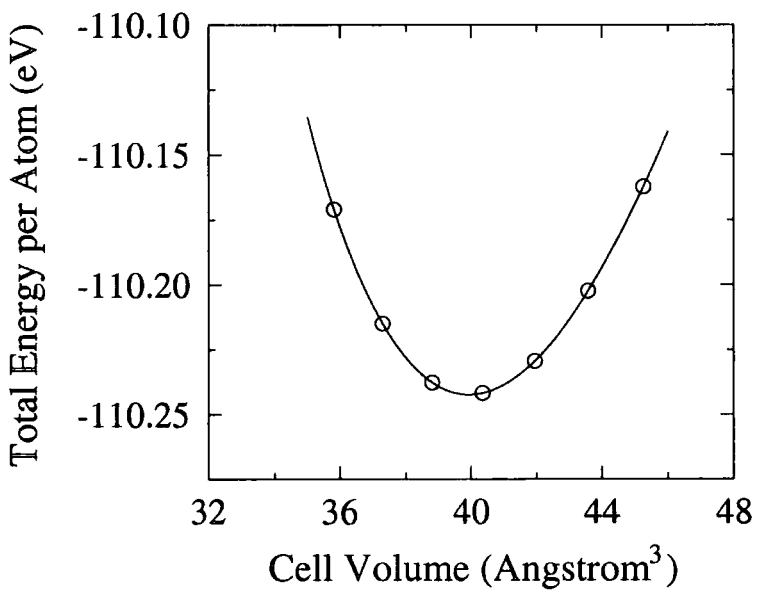

Figure 3.2: The open circles represent total energies calculated using the HCTH functional for different unit cell volumes for Si. The solid line shows the Murnaghan fit determined from relation (3.22), from which the equilibrium cell volume V_0 and bulk modulus B_0 are obtained.

Table 3.1 presents the calculated lattice constants obtained with the LDA, PW91 and HCTH functionals. As expected, the LDA uniformly underestimates with a mean error of −0.06 Å, while PW91 uniformly overestimates

them, with a mean error of +0.03 Å. HCTH performs worse than even the LDA with a systematic overestimation of +0.08 Å. For HCTH there is a clear correlation between the accuracy of the lattice constant and the number of occurrences of the constituent atoms in the G2 fitting data used to determine the functional. For C, Si and Ge, the lattice constant errors are −0.01 Å (∼ 0.3%), 0.07 Å (∼ 1.3%), and 0.14 Å (∼ 2.5%) respectively - the number of systems in the fitting data containing carbon, silicon, and germanium atoms are 19, 7, and 0 respectively. The errors for the aluminium semiconductors AlN, AlP, and AlAs are 0.06 Å (∼ 1.4%), 0.09 Å (∼ 1.7%), and 0.12 Å (∼ 2.1%) with 8, 4, and 0 occurrences of nitrogen, phosphorus, and arsenic atoms in the fitting data. The errors for the gallium semiconductors GaN, GaP, and GaAs are 0.07 Å (∼ 1.6%), 0.09 Å (∼ 1.7%), and 0.16 Å (∼ 2.8%) respectively. The fitting data contains one system with aluminium but none with gallium, which is consistent with the lattice constants for GaN and GaAs being less accurate than AlN and AlAs.

The significant overestimation of lattice constants made by HCTH is consistent with the findings of Kurth *et al.* [108]. Their unit cell volumes for Si, Ge, and GaAs correspond to lattice constants of 5.48, 5.80, and 5.80 Å respectively, which are close to the values shown in Table 3.1. It should be noted that the degradation in performance for heavier systems is also evident in molecular calculations [98]. A subset of the molecules in the HCTH fitting data have well-known experimental data. The subset includes 28 molecules containing just first-row atoms and 12 containing second-row atoms. HCTH overestimates the bond lengths of these two sets of systems by an average of 0.008 Å and 0.025 Å respectively [98].

3.3.2 Bulk Moduli

The bulk modulus B_0 of a crystal is defined as,

$$B_0 = \frac{1}{K} = -V\frac{dP}{dV} \tag{3.23}$$

Table 3.1: Optimised lattice constants (in Å) calculated using the LDA, PW91 and HCTH. Mean absolute errors (mae) from experiment are also given.

	LDA	PW91	HCTH	Expt.
C	3.53	3.57	3.56	3.57
Si	5.38	5.46	5.50	5.43
Ge	5.54	5.71	5.80	5.66
SiC	4.30	4.36	4.37	4.35
AlN	4.31	4.39	4.43	4.37[a]
AlP	5.41	5.49	5.54	5.45
AlAs	5.60	5.69	5.78	5.66
GaN	4.46	4.55	4.57	4.50[a]
GaP	5.38	5.49	5.54	5.45
GaAs	5.57	5.70	5.81	5.65
mae	0.06	0.03	0.08	–

[a]Reference [111]

where the pressure P is minus the derivative of the total energy,

$$P = -\frac{dE}{dV}\,. \tag{3.24}$$

Therefore B_0 effectively measures the curvature of the energy versus volume curve about the relaxed volume, V_0. Again, this can be computed from the same fit to the Murnaghan equation (3.22) used to determine a_0. Experimental bulk moduli for cubic crystals are determined from elastic constants [112] C_{11} and C_{12} using the formula [113]:

$$B_0 = (C_{11} + 2C_{12})/3\,. \tag{3.25}$$

Table 3.2 presents the calculated bulk moduli using the LDA, PW91 and HCTH for each system along with available experimental data. The HCTH values agree well with those of Kurth *et al.* [108]. The LDA and PW91 values are also in good agreement with other previous calculations [114, 115]. The

Table 3.2: Bulk moduli (in GPa) calculated using the LDA, PW91 and HCTH.

	LDA	PW91	HCTH	Expt.
C	457	425	428	442
Si	97	88	82	98.8
Ge	78	62	54	76.8
SiC	227	215	212	–
AlN	206	192	186	202[a]
AlP	89	82	78	–
AlAs	75	71	60	–
GaN	199	173	161	190[a]
GaP	89	77	69	88.7
GaAs	75	65	51	74.8
mae	5	13	20	–

[a]Reference [116]

LDA typically gives bulk moduli closer to experiment than both of the GGAs - the mean absolute errors from experiment are 5, 13 and 20 GPa for the LDA, PW91 and HCTH respectively. Whereas LDA tends to overestimate experiment, PW91 and HCTH consistently underestimate experiment.

3.3.3 Cohesive Energies

The cohesive energy of a solid is the energy required to break the atoms of the solid into isolated atomic species, i.e,

$$E_{\text{coh}} = E_{\text{solid}} - \sum_{\text{A}} E_{\text{A}}^{\text{isolated}} \tag{3.26}$$

where A represents the different atoms that constitute the solid. The cohesive energy is important, not only because it probes the energetic quality of a functional, but also because of the range of densities tested, namely infinite and finite corresponding to the solid and atomic systems respectively. Calculated values of the cohesive energy are compared with experimental results

which can be obtained by measuring the latent heat of sublimation at various low temperatures, and extrapolating to zero Kelvin.

The calculated cohesive energies are presented in Table 3.3. For the atomic calculations spin-dependent forms of all three functionals are employed, with the atoms in their ground-state spin configurations. The energy associated with the bulk solid is evaluated at the optimised lattice constant given in Table 3.1. Convergence tests show that a 10 Å supercell is sufficiently large to converge the total energy of each atom to better than 1 meV/atom.

Table 3.3: Cohesive energies (in eV/atom) calculated using the LDA, PW91 and HCTH.

	LDA	PW91	HCTH	Expt.
C	8.83	7.56	7.17	7.37
Si	5.29	4.56	4.12	4.63
Ge	4.67	3.73	3.23	3.85
SiC	14.78	12.76	11.79	–
AlN	13.17	11.28	10.08	–
AlP	9.67	8.29	7.33	–
AlAs	8.89	7.25	6.45	–
GaN	10.52	8.46	7.31	8.96[a]
GaP	8.67	7.02	5.98	–
GaAs	8.02	6.18	5.27	6.52
mae	1.20	0.24	0.85	–

[a]Reference [117]

Again, the LDA and PW91 values are in good agreement with the calculations reported in Refs. [114, 115, 118, 119]. The serious overbinding of LDA is clearly evident. While PW91 and HCTH go someway to correcting this overbinding, HCTH overcompensates, giving cohesive energies that are systematically lower than experiment. As with the lattice constants, the HCTH error increases as the periodic table is descended, from 0.20 eV (3%) in C to 1.25 eV (19%) in GaAs. The HCTH underbinding is consistent with

the overestimated lattice constants in Table 3.1 and the underestimated bulk moduli in Table 3.2.

3.3.4 Electronic Bandstructures

The electronic bandstructure of a solid shows the eigenvalues associated with the valence and conduction bands along specific directions in the Brillouin zone of that particular crystal structure. One of the most important reasons for computing bandstructures is to determine the band gap, i.e. the difference between the highest valence band and the lowest conduction band energies, since this can provide insight into potentially useful materials for optical device applications. However band gaps calculated from Kohn-Sham eigenvalues using the LDA and the GGA notoriously underestimate the results obtained from experiment. In fact this is one property where the GGA provides no consistent improvement over the LDA, and is sometimes worse. Discrepancies between DFT and experiment are in the range 50% to 100%.

Problems arise in calculating band gaps because the Kohn-Sham eigenvalues do not have a strict physical interpretation (except for the highest occupied level [120, 121]) and also because the exact exchange-correlation potential exhibits a discontinuity when the number of electrons in the system passes through an integer [121, 122, 123], which is not described by the usual *continuum* description [37] provided by the LDA and the GGA. The size of the discontinuity therefore sets a limit on the accuracy that can be achieved with continuum functionals, although its exact value in solids is unknown. It is generally believed to be a significant proportion of the actual gap [124, 125], however Städele *et al.* [126] recently suggested that it may be smaller than originally thought. From a purely pragmatic viewpoint, it would be highly advantageous to develop functionals that improve upon the particularly poor description yielded at present, and so give band gaps in reasonable agreement with experiment.

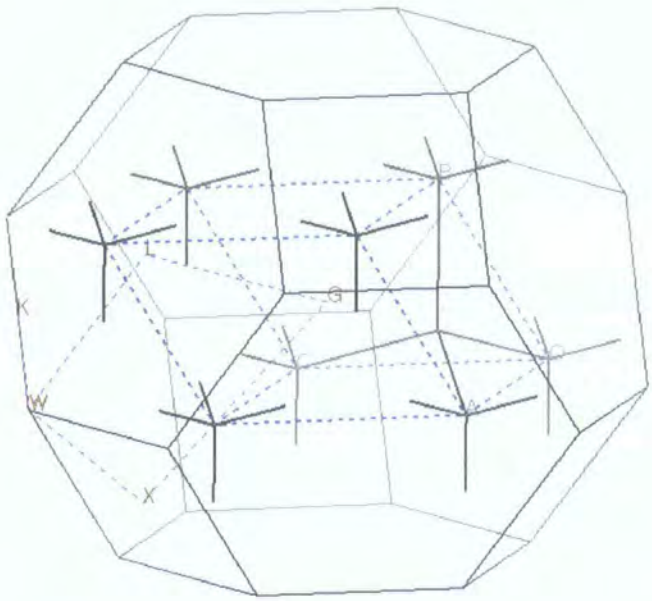

Figure 3.3: The primitive cell (purple dashed lines) and corresponding first Brillouin zone (black lines) for the diamond/zinc-blende structure. The positions of high symmetry in the Brillouin zone are indicated by the brown characters - G represents the Γ point, and the k-point path used in the bandstructure calculations is also shown by the blue dashed lines.

To determine the bandstructure of a material, a single-point energy calculation is first performed at a specified crystal geometry to obtain the self-consistent groundstate density. This fixes the form of the Kohn-Sham Hamiltonian which is then solved to give the corresponding Kohn-Sham eigenvalues. The eigenvalues are computed at a greater number of k-points, along specific directions in the Brillouin zone, than the ones used in the single-point energy calculation. Fig. 3.3 shows the Brillouin zone associated with the diamond/zinc-blende structure, along with the k-point path used in the bandstructure calculations performed here.

Table 3.4 displays minimum band gaps calculated at optimised lattice constants; values calculated at experimental lattice constants are given in

Table 3.4: Minimum electronic band gaps (in eV) calculated using the LDA, PW91, and HCTH at optimised lattice constants. Values in parentheses are band gaps calculated at experimental lattice constants.

	LDA	PW91	HCTH	Expt.
C	4.21 (4.12)	4.20 (4.20)	4.24 (4.22)	5.48
Si	0.44 (0.49)	0.62 (0.60)	0.87 (0.80)	1.17
Ge	0.35 (0.03)	0.00 (0.18)	0.00 (0.32)	0.71
SiC	1.30 (1.33)	1.42 (1.42)	1.63 (1.63)	2.39
AlN	3.29 (3.26)	3.39 (3.39)	3.68 (3.70)	–
AlP	1.38 (1.42)	1.61 (1.56)	1.99 (1.89)	2.50
AlAs	1.26 (1.32)	1.42 (1.39)	1.81 (1.66)	2.32
GaN	1.90 (1.74)	1.51 (1.73)	1.68 (1.99)	3.45[a]
GaP	1.33 (1.40)	1.55 (1.55)	1.65 (1.80)	2.35
GaAs	0.64 (0.32)	0.30 (0.49)	0.13 (0.72)	1.52
mae	1.01 (1.08)	1.03 (0.97)	0.88 (0.76)	–

[a]Reference [127]

parentheses. At optimised lattice constants, the HCTH band gaps are larger than LDA and PW91 for all systems except Ge, GaN, and GaAs. At experimental lattice constants, the HCTH band gaps are generally larger than at the optimised lattice constants as a result of lattice contraction. The PW91 and HCTH bandstructures for Si are superimposed in Fig. 3.4. In general the main differences between the two GGAs occur for the conduction bands away from the Γ point, so although the gap increases with HCTH, the shift in energy is not uniform across the Brillouin zone.. The conduction band minimum (CBM) for Si correctly occurs at a point along the Γ-X axis in the Brillouin zone for all three functionals. The HCTH functional decreases the valence band width, i.e. the difference between the lowest highest valence band energies, with respect to PW91 by ~ 0.2 eV.

In Ge, the CBM and valence band maximum (VBM) touch at the Γ point

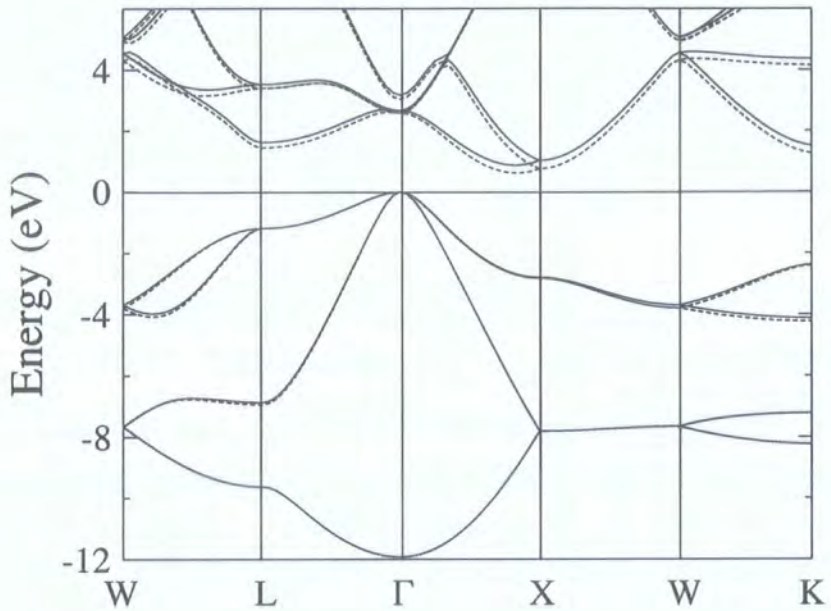

Figure 3.4: The bandstructure of bulk Si calculated using HCTH (solid lines) and PW91 (dashed lines). The top valence bands have been aligned at the Γ point (zero energy).

for both HCTH and PW91, therefore giving no gap. The systems SiC, AlP, and AlAs have indirect band gaps for all three functionals, with the CBM correctly occuring at the X point in all cases. Each functional also correctly yield direct gaps for GaAs, although are considerably underestimated - the GGAs give the worst agreement for this system.

Considering the III-V nitride semiconductors, each functional correctly predicts GaN to have a direct transition at the Γ point, although LDA gives a larger gap than both PW91 and HCTH at the optimised lattice constant. AlN has an indirect band gap occurring at the X-point with each functional, despite the direct nature found by experiment. The band gap at the Γ point for AlN is calculated to be 4.62, 4.13 and 4.13 eV for LDA, PW91 and HCTH respectively, at the optimised lattice constant.

3.3.5 Charge Density Differences

Molecular studies have revealed that the LDA and GGA give a reasonable description of the exchange-correlation potential in regions close to a molecule, but break down with increasing distance [37]. This is evident since the true potential decays asymptotically with an inverse distance behaviour, whereas the LDA and GGA potentials behave exponentially at large distance from a finite charge distribution. As a consequence, conventional functionals such as PW91 lead to over-diffuse electron densities. This also has an effect on the occupied-unoccupied eigenvalue differences in atoms and molecules [128], which are greatly underestimated.

The charge density difference of Si between HCTH and PW91 has been computed to investigate whether HCTH provides any improvement regarding the diffuse nature of the density. Fig. 3.5(a) shows the crystal lattice of Si, and the colour contours in Figs. 3.5(b) and (c) display the positive difference between the densities calculated as, PW91 - HCTH and HCTH - PW91 respectively. The density differences are superimposed onto the Si crystal lattice in order to illustrate the regions where charge has been transferred. It is clear from these plots that the HCTH density is greatest near the atomic sites and bonds, whereas the PW91 density is higher in the interstitial regions of the lattice. The superposition of Figs. 3.5(b) and (c) given in 3.5(d) clearly shows the transfer of density.

If the over-diffuse nature of the density in the molecular environment extends into the solid state, then the above results indicate that HCTH yields a more accurate description of the exchange-correlation potential than PW91. This would also explain why HCTH band gaps are generally more accurate than those obtained with PW91.

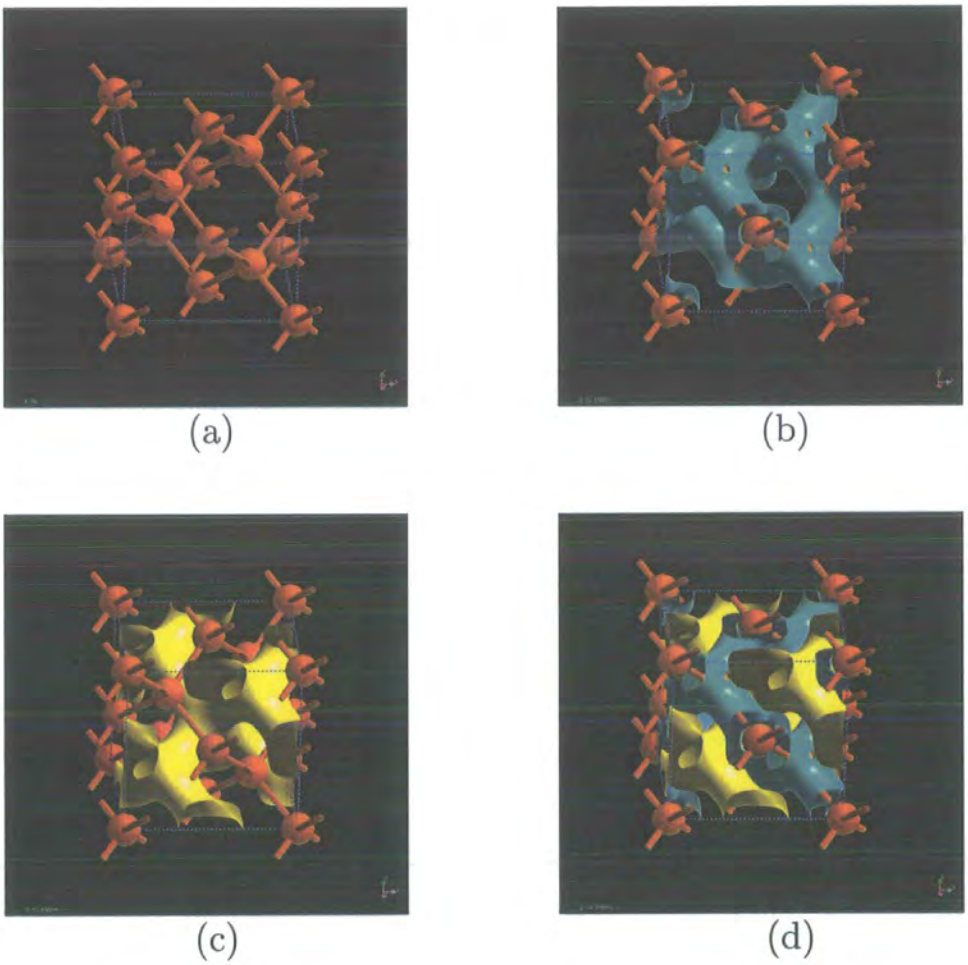

Figure 3.5: The crystal lattice of Si is displayed in (a) while the positive contributions to the density differences are shown by the contours, for PW91 minus HCTH (b) and HCTH minus PW91 (c). The superposition of (b) and (c) is shown in (d) to illustrate the net flow of charge from the interstitials to the bonding regions in going from PW91 to HCTH.

Table 3.5: HCTH lattice constants, a_0, and cohesive energies, E_{coh}, calculated using LDA pseudopotentials. Comparisons should be made with the consistent pseudopotential results presented in Tables 3.1 and 3.3.

	a_0(Å)	E_{coh}(eV)
C	3.51	7.36
Si	5.41	4.08
Ge	5.68	3.27
SiC	4.32	11.36
GaN	4.34	9.61
GaP	5.41	6.78
GaAs	5.65	6.38

3.3.6 Pseudopotential Considerations

A comment is made on the influence of pseudopotentials on computed properties. The results presented previously were obtained using consistent pseudopotentials calculated for each of the three functionals studied. Calculations were also performed using the GGAs inconsistently with LDA pseudopotentials. In general, the weakening of the bonding with respect to the LDA - that is a characteristic feature of the GGA - was not observed with HCTH nor PW91. HCTH lattice constants were often significantly shorter, mostly lying below the experimental values, and cohesive energies were greater than in the consistent case. Examples of inconsistent results obtained with HCTH are shown in Table 3.5. It is clear that an inconsistent treatment of exchange-correlation interactions for the core and valence electrons leads to false assessment of the functional being tested. The same findings were reached in other pseudopotential studies [115].

3.4 Performance of Other HCTH-Type Functionals

This section presents a more complete assessment of the semi-empirical functional concept by examining variants of the HCTH functional in which the 15 coefficients are obtained by different fit strategies. The fitting procedures for the new HCTH functionals were carried-out by Wilson *et al.* [129]. Brief descriptions of the functionals are presented in Table 3.6, and the expansion coefficients are given in Tables B.1 and B.2.

Table 3.6: Description of the five variants of the HCTH-GGA, whereby the 15 expansion parameters are fitted to different sets of data. The acronym "HEG" indicates the homogeneous electron gas.

Functional	Description
1. HCTH - HEG	Enforcing the HEG condition
2. HCTH - 26	Fitting to a restricted set of 26 systems
3. HCTH - $\frac{1}{4}$	Fitting to *ab-initio* potentials only
4. HCTH - 26 - $\frac{1}{4}$	Combination of (2) and (3)
5. HCTH - HEG - 26 - $\frac{1}{4}$	Combination of (1), (2) and (3)

The first issue to investigate is the effect of explicitly enforcing the homogeneous electron gas constraint (HEG) which is traditionally considered to be an important ingredient for functionals applied to solid state calculations [54], in contrast to molecular calculations where it is not considered so important [130]. Recently, Kurth *et al.* [108] observed a trend between the accuracy of equilibrium cell volumes (lattice constants) and functionals that satisfied the HEG condition. The HCTH-HEG functional uses exactly the same fit data as the original HCTH, except only 12 coefficients are optimised - the first three are set to 1 in order to regain the LDA in the HEG

limit. The results from the previous section also reveal that the quality of the semiconductor properties yielded by HCTH are dictated by the occurrence of the atoms used in the fit data. The HCTH-26 functional will be used to examine this trend since it employs a restricted data-set of systems drawn uniformly across the first two rows of the periodic table, in contrast to HCTH which is biased toward lighter systems. Included within this set are systems such as N_2 and P_2. HCTH-$\frac{1}{4}$ was constructed using only *ab-initio* exchange-correlation potentials and has so far proved very successful for structural properties of molecules [131]. Interestingly this functional provides uniform accuracy across the first periodic table when applied to molecules [131]. The two remaining functionals, HCTH-26-$\frac{1}{4}$ and HCTH-HEG-26-$\frac{1}{4}$, were determined using a combination of the previous approaches.

3.4.1 Computational Results

The calculations were performed on five systems: C, Si, Ge, GaAs and GaN, using consistently generated pseudopotentials. Table 3.7 shows the mean absolute error for the structural, cohesive and electronic properties obtained using these functionals. Corresponding LDA, PW91 and HCTH values are also given for comparison.

The HCTH-HEG functional provides a definite improvement upon HCTH for all quantities examined which re-affirms the importance of the HEG condition in solids. This functional however is not as accurate as PW91 - for example the average error in the cohesive energy is over double that of PW91. It should be noted that HCTH-HEG provides a worse description than HCTH for molecules [132]. The HCTH-26 functional was designed to reduce the bias from the fitting set, however as can be seen from Table 3.8 it displays the same trend as HCTH on moving to systems lower down in the periodic table, although it is not as pronounced as with HCTH.

HCTH-$\frac{1}{4}$ is the most successful of all the semi-empirical functionals tested.

Table 3.7: Mean absolute errors (mae) from experiment for other HCTH-type functionals for the lattice constant (in Å), bulk modulus (in GPa), cohesive energy (in eV) and electronic band gap (in eV) of the bulk systems C, Si, Ge, GaN and GaAs.

Functional	a_0	B_0	$E_{\rm coh}$	$E_{\rm gap}$
LDA	0.07	5	1.20	0.96
PW91	0.04	14	0.24	1.14
HCTH	0.08	23	0.85	1.08
HCTH-HEG	0.06	14	0.56	1.05
HCTH-26	0.07	15	0.67	1.04
HCTH-$\frac{1}{4}$	0.02	11	0.56	0.97
HCTH-26-$\frac{1}{4}$	0.08	20	1.27	1.09
HCTH-HEG-26-$\frac{1}{4}$	0.11	26	0.78	1.17

The mean absolute error in lattice constants is halved in comparison to PW91, and bulk moduli are also closer to experiment than PW91, although the LDA is still the most accurate for this quantity. As can be seen from Table 3.8 it provides excellent agreement with experiment for the heavier atomic systems, GaAs and GaN. HCTH-$\frac{1}{4}$ also provides the best description of band-gaps of the all the GGAs tested, although it is not as good as the LDA for the systems considered. As is the case with molecules, the energetic predictions of this functional are not particularly successful - although the cohesive energies are an improvement over HCTH, they are still significantly worse than PW91.

The two functionals that use a mixture of the fitting procedures give results that are generally worse than HCTH.

The interesting characteristics of the HCTH-$\frac{1}{4}$ functional - excellent structural properties for solids and molecules, uniform accuracy despite the biased fitting set, but poor energetics - are reflected by its distinctive enhancement factor $F_{\rm XC}(r_s, s)$, which is shown in Fig. 3.6. The enhancement factor of

Table 3.8: Absolute percentage errors between the calculated and experimental lattice constants and bulk moduli for five semiconductor systems.

	%Error in a_0			%Error in B_0		
	HCTH	HCTH-26	HCTH-$\frac{1}{4}$	HCTH	HCTH-26	HCTH-$\frac{1}{4}$
C	0.28	0.42	1.15	3	2	3
Si	1.29	1.00	0.09	17	11	9
Ge	2.47	1.69	0.32	30	21	14
GaN	1.56	1.31	0.28	15	13	6
GaAs	2.83	2.79	0.47	32	33	16

HCTH and PW91 are also shown for comparison. It is very clear from these plots that the three functionals are very different. Since the semi-empirical fitting process allows limited control over the form of the resulting functional, beyond that of the choice of the systems employed in the fit data, exact conditions are often violated. Indeed important conditions such as the sum rules on the exchange and correlation holes will almost certainly not be satisfied by semi-empirical functionals. It can be observed from Fig. 3.6, that HCTH dramatically violates two important constraints on $F_{\rm XC}(r_s, s)$, namely the Lieb-Oxford bound [57],

$$F_{\rm XC}(r_s, s) \leq 2.27\,, \tag{3.27}$$

and the non-crossing condition,

$$F_{\rm XC}(r_s', s) > F_{\rm XC}(r_s, s); \quad (r_s' > r_s)\,, \tag{3.28}$$

which arises from the scaling behaviour of the correlation energy [56]. Interestingly, HCTH-$\frac{1}{4}$ violates the Lieb-Oxford bound in a greater way than HCTH, as observed by the the lines corresponding to high r_s values, although it is clear that the non-crossing condition is satisfied. The degradation in the accuracy of the cohesive energies directly correlates to the amount by which the $F_{\rm XC}(r_s, s) = 2.27$ bound is violated by HCTH and HCTH-$\frac{1}{4}$. This

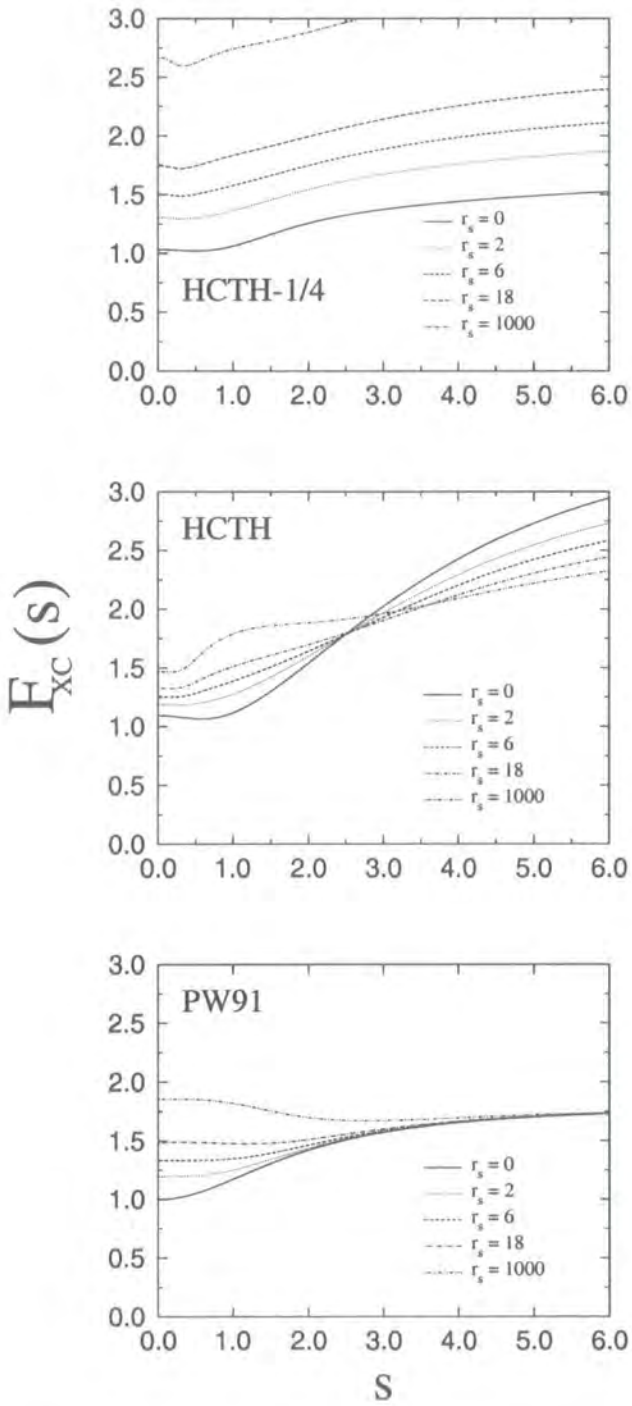

Figure 3.6: The GGA enhancement factor, $F_{\rm XC}(r_s, s)$, for HCTH-$\frac{1}{4}$, HCTH and PW91 at various r_s values.

indicates that the Lieb-Oxford condition is important for the successful determination of cohesive energies - at least for the type of systems studied here.

Attempts have been made by Menconi and Tozer [133] to improve the energetic properties of molecules with the HCTH-$\frac{1}{4}$ functional. They find that this can only be achieved at the expense of a subsequent degradation in the quality of structural predictions.

3.5 Conclusions

The principle aim of the work in this chapter was to assess the suitability of employing semi-empirical GGAs in solid state calculations. Due to a proven success in molecular applications, and a highly flexible functional form, the HCTH functional was used to calculate structural, cohesive and electronic properties of several semiconductors, and its performance was judged in relation to the LDA and the non-empirical PW91. HCTH did improve upon the LDA and PW91 for electronic band gaps, however it gave the worst structural properties of the three functionals; cohesive energies were also much worse than PW91. None of the other HCTH variant functionals were able to consistently improve upon PW91 for all four quantities studied. Small improvements can be made such as enforcing the homogeneous electron gas condition, although this is still not enough to succeed over PW91.

GGAs that satisfy the homogeneous gas constraint are known to cause a degradation in molecular properties, consequently this is often deemed unimportant in the chemistry community and is frequently ignored in molecular derived functionals. However it is evident that this condition is a necessary ingredient for the successful description of solids, as demonstrated here, so a universally applicable functional must incorporate this condition. This issue is a simple illustration of the inflexibility of the GGA form - an improvement in one particular property or type of system (e.g. molecular or solid state) is

often at the expense of another property or system.

Two important conclusions can be deduced from this study. Firstly, semi-empirical fitting procedures are not an appropriate way to determine universally applicable functionals, since the properties of the functional are biased by the data in the fitting set - which at the moment can only be obtained for a limited number of systems in the upper rows of the periodic table. Secondly, although the GGA has proved to be a success over the LDA, it is a restricted form that has probably reached the limit of its capabilities. For these reasons it is necessary to pursue different functional forms that supersede the limited GGA, and that are constructed using sound theoretical principles rather than semi-empirical fits. The remainder of this thesis is dedicated to one such functional - the weighted density approximation.

Chapter 4

The Weighted Density Approximation

4.1 Introduction

Motivated by the conclusions of the Chapter 3, and the considerations made at the end of Chapter 2 regarding existing exchange-correlation functionals, the weighted density approximation (WDA) [85, 86, 87] is now investigated as a possible route toward more sophisticated and accurate functionals. The WDA is a physically satisfying choice simply because of its affinity with the exact functional form given by the adiabatic connection method [134]. Consequently, the WDA is a truly non-local functional of the density, which is at variance with all other post-GGA treatments being actively investigated at the moment such as the MGGA [73].

Although the fully non-local character of the WDA brings various physical advantages, as shown in Table 2.1, the WDA is more difficult to implement in a tractable computational scheme than the much simpler LDA or GGA approaches. Consequently, the development of the WDA effectively came to a halt soon after its inception in the 1970s, due to Perdew-Zungers parametrisation of the LDA [50] in 1981, and the advent of the GGA later-

on. Since then, the WDA has remained in relative obscurity, seldom used or investigated. Nevertheless, the WDA has been cast into a computationally tractable scheme [135, 136] based on a reciprocal-space formalism, which although will never compete with the speed of semi-local functionals, does allow substantial savings in computational expenditure compared with the direct real-space approach. This chapter provides details of the reciprocal-space implementation of the WDA that is particularly suited to a periodic representation of the density.

4.2 Theory of the WDA

4.2.1 The Functional Form

As with all types of functional, the manner in which the (coupling-constant averaged) exchange-correlation (XC) hole is modelled characterises the functional form. To recap, the exact hole $n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')$ has the general form given by relation (2.24),

$$n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}') = n(\mathbf{r}')[g_{\mathrm{XC}}(\mathbf{r},\mathbf{r}') - 1]\,. \tag{4.1}$$

In the LDA the, non-local dependence on the density, $n(\mathbf{r}')$, is replaced by the local density $n(\mathbf{r})$, and the pair-correlation function is that obtained from the homogeneous electron gas $g_{\mathrm{XC}}^{\mathrm{hom}}(r)$, which uses the modulus of the distance $r = |\,\mathbf{r}-\mathbf{r}'\,|$,

$$n_{\mathrm{XC}}^{\mathrm{LDA}}(\mathbf{r},\mathbf{r}') = n(\mathbf{r})[g_{\mathrm{XC}}^{\mathrm{hom}}(r) - 1]\,. \tag{4.2}$$

The local density dependence in (4.2) means that for finite inhomogeneous systems such as atoms and molecules, the LDA prescription will lead to an overestimate of the true XC hole value at points of high density, and an underestimate at low densities. However, the LDA in general yields sensible energies for many systems as result of the partial cancellation of these errors in the XC hole [87]. The LDA hole also obeys the fundamental sum rule given

in (2.8) that constrains the XC energy to sensible bounds. In an attempt to improve upon the LDA description, Gunnarsson *et al.* proposed the average density approximation (ADA) [84], whereby $n(\mathbf{r}')$ is replaced by a density average $\bar{n}(\mathbf{r})$ over the extent of the hole, rather than just $n(\mathbf{r})$. The ADA ansatz is:

$$n_{\mathrm{XC}}^{\mathrm{ADA}}(\mathbf{r},\mathbf{r}') = \bar{n}(\mathbf{r})[\tilde{g}_{\mathrm{XC}}^{\mathrm{hom}}(r,\bar{n}(\mathbf{r})) - 1] \tag{4.3}$$

where $\tilde{g}_{\mathrm{XC}}^{\mathrm{hom}}(r,\bar{n}(\mathbf{r}))$ is the homogeneous electron gas form, but the averaged density $\bar{n}(\mathbf{r})$ is used instead of the actual density. The average density is determined by,

$$\bar{n}(\mathbf{r}) = \int W[r,\bar{n}(\mathbf{r})]\, n(\mathbf{r}')\, d\mathbf{r}' \,, \tag{4.4}$$

where the weight function $W[r,\bar{n}(\mathbf{r})]$ must satisfy the normalisation condition,

$$\int W[r,\bar{n}(\mathbf{r})]\, d\mathbf{r}' = 1 \,. \tag{4.5}$$

The ADA however has certain deficiencies and the error cancellation is usually not as complete as in the LDA [87].

In the same paper that the ADA was proposed, Gunnarsson *et al.* devised the weighted density approximation, which was also arrived at independently by Alonso and Girifalco [85]. The WDA hole retains the same non-local density dependence as the exact result given in (2.24), and is usually modelled in terms of a simple analytic function, $G^{\mathrm{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r}))]$,

$$n_{\mathrm{XC}}^{\mathrm{WDA}}(\mathbf{r},\mathbf{r}') = n(\mathbf{r}')\, G^{\mathrm{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})] \,. \tag{4.6}$$

The specific choice of $G^{\mathrm{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$ is investigated in this work, suffice it to say that it must obey a minimum number of exact limiting conditions, which will be elaborated on in Sec. 4.2.3, in order to be physically sensible. The quantity $\tilde{n}(\mathbf{r})$ is a non-local parameter called the weighted density, and

is determined at each point in space by satisfying the sum rule relation (2.8) on $n_{\rm XC}^{\rm WDA}(\mathbf{r},\mathbf{r}')$,

$$\int n(\mathbf{r}')\, G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]\, d\mathbf{r}' = -1\,. \tag{4.7}$$

Substituting (4.6) into the exact formula for $E_{\rm XC}[n(\mathbf{r})]$, given by (2.9), leads to the definition of the WDA energy functional,

$$E_{\rm XC}^{\rm WDA}[n(\mathbf{r})] = \frac{1}{2}\int n(\mathbf{r})\, d\mathbf{r} \int \frac{n_{\rm XC}^{\rm WDA}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}'\,. \tag{4.8}$$

Form (2.11), the corresponding WDA energy-density $\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ is,

$$\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r}) = \frac{1}{2}\int \frac{n_{\rm XC}^{\rm WDA}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}'\,. \tag{4.9}$$

The WDA is a conceptually simple XC functional that is based on the exact expression given by the adiabatic connection method. The WDA essentially models the exact, but unknown, XC hole by retaining the correct nonlocal dependence on the density and through the use of analytic expressions, $G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$. The form of the WDA functional is therefore immediately understandable - this should be contrasted with the complicated GGA forms given in the previous chapters.

4.2.2 The WDA Potential

The exchange-correlation potential of any density-based functional is obtained by taking the functional derivative of the energy, $E_{\rm XC}[n(\mathbf{r})]$, with respect to $n(\mathbf{r})$, as given by relation (1.50). For the WDA this leads to three terms,

$$v_{\rm XC}^{\rm WDA}(\mathbf{r}) = \frac{\delta E_{\rm XC}^{\rm WDA}[n(\mathbf{r})]}{\delta n(\mathbf{r})} = v_1(\mathbf{r}) + v_2(\mathbf{r}) + v_3(\mathbf{r})\,. \tag{4.10}$$

The first term is simply the energy density $\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ given by (4.9),

$$v_1(\mathbf{r}) = \varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r}) = \frac{1}{2}\int n(\mathbf{r}')\, \frac{G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}'\,. \tag{4.11}$$

The second term, $v_2(\mathbf{r})$, arises from the asymmetry of $G^{\text{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$ with respect to the interchange of particle coordinates $\mathbf{r}$ and $\mathbf{r}'$, which ultimately orginates from the symmetry of the pair-density $P(\mathbf{r},\mathbf{r}')$ given by (2.13),

$$v_2(\mathbf{r}) = \frac{1}{2}\int n(\mathbf{r}')\,\frac{G^{\text{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r}')]}{|\,\mathbf{r}-\mathbf{r}'\,|}\,d\mathbf{r}'\,. \tag{4.12}$$

Finally, $v_3(\mathbf{r})$ accounts for the dependence of $\tilde{n}(\mathbf{r})$ on the density $n(\mathbf{r})$, through $G^{\text{WDA}}[\mathbf{r}',\mathbf{r}'';\tilde{n}(\mathbf{r}')]$,

$$v_3(\mathbf{r}) = \frac{1}{2}\int n(\mathbf{r}')\,d\mathbf{r}'\int \frac{n(\mathbf{r}'')}{|\,\mathbf{r}'-\mathbf{r}''\,|}\times\frac{\delta G^{\text{WDA}}[\mathbf{r}',\mathbf{r}'';\tilde{n}(\mathbf{r}')]}{\delta n(\mathbf{r})}\,d\mathbf{r}''\,. \tag{4.13}$$

The functional derivative of $G^{\text{WDA}}[\mathbf{r}',\mathbf{r}'';\tilde{n}(\mathbf{r}')]$ in (4.13) can be re-written using the chain rule for functional differentiation as,

$$\frac{\delta G^{\text{WDA}}[\mathbf{r}',\mathbf{r}'';\tilde{n}(\mathbf{r}')]}{\delta n(\mathbf{r})} = \frac{\partial G^{\text{WDA}}[\mathbf{r}',\mathbf{r}'';\tilde{n}(\mathbf{r}')]}{\partial\tilde{n}(\mathbf{r}')}\times\frac{\delta\tilde{n}(\mathbf{r}')}{\delta n(\mathbf{r})}\,, \tag{4.14}$$

and since the weighted density is related to the actual density through the sum rule (4.7), it is possible to write the derivative of the weighted density as [137]:

$$\frac{\delta\tilde{n}(\mathbf{r}')}{\delta n(\mathbf{r})} = -\frac{G^{\text{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r}')\,]}{\displaystyle\int n(\mathbf{r}'')\,\partial G^{\text{WDA}}[\mathbf{r}',\mathbf{r}'';\tilde{n}(\mathbf{r}')]/\partial\tilde{n}(\mathbf{r}')\,d\mathbf{r}''}\,. \tag{4.15}$$

Evaluating $v_2(\mathbf{r})$ and $v_3(\mathbf{r})$ is computationally demanding compared with calculating $\varepsilon_{\text{XC}}^{\text{WDA}}(\mathbf{r})$ and the sum rule, consequently early WDA studies [135] made the approximation $v_1(\mathbf{r}) = v_2(\mathbf{r})$, and also ignored $v_3(\mathbf{r})$. However all three terms are required for a complete self-consistent potential, so the full WDA potential is implemented in the present study.

4.2.3 Model Pair-Correlation Functions

The form of the model function, $G^{\text{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$, is the most important quantity in the WDA since it contains all of the approximations in the theory. It is related to the pair-correlation function $g_{\text{XC}}(\mathbf{r},\mathbf{r}')$, by a simple relation,

$$G^{\text{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})] = g_{\text{XC}}(\mathbf{r},\mathbf{r}') - 1\,, \tag{4.16}$$

however, due to the trivial connection between the two quantities in (4.16), $G^{\mathrm{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$ will also be referred to as the pair-correlation function. As a first model for $G^{\mathrm{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$, Gunnarsson and Jones used the exact exchange expression for the homogeneous electron gas, derived from the exchange energy in Hartree-Fock theory (see Parr and Yang [12]),

$$g_{\mathrm{X}}^{\mathrm{hom}}(y) = -\frac{9}{2}\left(\frac{\sin y - y\cos y}{y^3}\right)^2 , \tag{4.17}$$

where $y = |\,\mathbf{r}-\mathbf{r}'\,|\,k_{\mathrm{F}}(\tilde{n})$, and the Fermi wavevector depends on the weighted density, not the local density $n(\mathbf{r})$,

$$k_{\mathrm{F}}(\tilde{n}) = [3\pi^2\,\tilde{n}(\mathbf{r})]^{\frac{1}{3}} . \tag{4.18}$$

This model, which leads to the Dirac expression for the homogeneous gas exchange energy given by (1.16), was reasonably successful in the calculation of atomic exchange energies, when used in the WDA [10]. However correlation effects are obviously not included within this pair-correlation function. Guided by exact limiting conditions, Gunnarsson and Jones [138] proposed a modified WDA scheme that describes exchange and correlation together, in which an extremely simple analytic model, $G^{\mathrm{GJ}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$, was chosen that depends on two non-empirically defined parameters, $\alpha(\tilde{n})$ and $\beta(\tilde{n})$,

$$G^{\mathrm{GJ}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})] = \alpha(\tilde{n})[1-\mathrm{e}^{-(\beta(\tilde{n})/|\mathbf{r}-\mathbf{r}'|)^5}] . \tag{4.19}$$

The parameters are determined from two exact conditions: that the XC energy density and sum rule are satisfied for a homogeneous electron gas with the local density value $n(\mathbf{r})$. As a result, this prescription also correctly reduces to the LDA in the limit of a homogeneous electron gas. The particular exponential form was chosen as it reproduces the correct asymptotic form for the XC potential outside the surface of a metal. However there is no fundamental reason for choosing this precise form. A more general expression for this type of function can be written using the substitution, $u = |\,\mathbf{r}-\mathbf{r}'\,|/\beta(\tilde{n})$:

$$G^{\mathrm{WDA}}[\,u,\tilde{n}(\mathbf{r})\,] = \alpha(\tilde{n})f(u) , \tag{4.20}$$

where, in the case of (4.19) the function $f(u)$ is given by,

$$f(u) = (1 - \mathrm{e}^{-u^{-5}}) . \tag{4.21}$$

Pair-correlation functions of the generalised form (4.20) will be primarily investigated in this work and various $f(u)$ will be used, guided by the few known constraints on the exact pair-correlation function. The simplification made in (4.19) is important since it facilitates the calculation of the functional derivative $\delta G^{\mathrm{WDA}}/\delta n(\mathbf{r})$ that is necessary for a self-consistent XC potential in (4.13). However, it also leads to an incorrect characteristic of the WDA pair-correlation functions in that they are not symmetric with respect to interchanging particle coordinates, $\mathbf{r}$ and $\mathbf{r}'$, i.e.

$$G^{\mathrm{WDA}}[\mathbf{r}, \mathbf{r}'; \tilde{n}(\mathbf{r})] \neq G^{\mathrm{WDA}}[\mathbf{r}', \mathbf{r}; \tilde{n}(\mathbf{r})] . \tag{4.22}$$

Nonetheless, this deficiency has been shown to have little influence on calculated properties [139], especially at the level of accuracy of present models

4.2.4 Self-Interaction

A crucial aspect in which all local and semi-local functionals fail is the complete cancellation of the Hartree self-repulsion energy for single-electron systems such as the hydrogen atom, as discussed in Sec. 2.1.5. This gives rise to a self-interaction error which plagues these functionals. The WDA on the other hand provides the correct description of one-electron systems, in principle. Complete self-interaction correction is not attained in practice because of the approximate nature of model pair-correlation functions such as the one proposed by Gunnarsson and Jones in (4.19). For a one-electron system $g_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ must be zero at all points since there is no chance of finding another electron in the system, and for this to be obtained in the WDA, the weighted density must be identically zero at all $\mathbf{r}$, which is not achieved precisely in actual calculations.

4.2.5 Asymptotic Properties

4.2.5.1 XC Energy Density

For a localised density distribution such as an atom, when $r = |\mathbf{r}-\mathbf{r}'|$ is large and $\mathbf{r} \to \infty$, the $\mathbf{r}'$ dependence in the factor $1/|\mathbf{r}-\mathbf{r}'|$ diminishes. It follows from this fact, together with the sum rule condition, that the asymptotic form of the WDA energy density (4.9) behaves as,

$$\lim_{r\to\infty} \varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r}) \to -\frac{1}{2\,r} \tag{4.23}$$

which is the exact result [87].

4.2.5.2 XC Potential

The asymptotic behaviour of the WDA potential is however slightly different from the exact result. The final term in the WDA potential containing the derivative of the pair-correlation function decays most rapidly, and $v_2(\mathbf{r})$ decays faster than $v_1(\mathbf{r})$, so the asymptotic behaviour of the WDA potential is dictated by $v_1(\mathbf{r})$. Using the same arguments as those used to obtain (4.23) leads to,

$$\lim_{r\to\infty} v_{\rm XC}^{\rm WDA}(\mathbf{r}) \to -\frac{1}{2\,r}\,, \tag{4.24}$$

which differs from the exact result by the prefactor of $1/2$. A model pair-correlation function satisfying the correct symmetry property would lead to, $v_1(\mathbf{r}) = v_2(\mathbf{r})$, and therefore,

$$v_{\rm XC}^{\rm WDA}(\mathbf{r}) = 2v_1(\mathbf{r}) + v_3(\mathbf{r}) \tag{4.25}$$

which reproduces the correct $-1/r$ limit.

The asymptotic limits of the WDA energy density and potential considerably improve upon local and semi-local functionals which decay exponentially to zero at large distance.

4.3 Computational Implementation

Although all of the WDA equations described previously have been stated in real-space form, practical calculations of the WDA are most efficiently performed within a reciprocal-space approach. Consequently, the WDA is most suited to solid-state applications using the plane-wave pseudopotential formalism. The first such implementation was given by Kerker [135] and was later expanded upon by Hybertsen and Louie [136], who derived the reciprocal-space equations for the complete WDA potential. The reformulation of the WDA sum rule, energy and potential, in terms of the reciprocal-space formalism are now described, along with details of an expedient implementation within the periodic CASTEP code [36].

4.3.1 Reciprocal-Space Representation

An efficient way of calculating the integral equations for the sum rule, total energy and potential in the WDA is to make use of the convolution theorem within a periodic representation of the density [140]. The convolution of two functions $a(\mathbf{r})$ and $b(\mathbf{r})$ can be written as,

$$\int a(\mathbf{r}')\, b(\mathbf{r}-\mathbf{r}')\, d\mathbf{r}' = \sum_{\mathbf{G}} A(\mathbf{G})B(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} \tag{4.26}$$

where $A(\mathbf{G})$ and $B(\mathbf{G})$ are Fourier transforms of $a(\mathbf{r})$ and $b(\mathbf{r})$ respectively, and the summation is over reciprocal-space lattice vectors, $\mathbf{G}$. If the functions are periodic and relatively smooth then this amounts to a substantial reduction in computational cost since the expensive real-space integrals are converted to multiplications in reciprocal-space. Indeed these equations would be in exactly the correct form if the weighted denisty were constant throughout all space, however this is not the case for inhomogeneous densities. Therefore in order to exploit this method it is necessary to discretise space and calculate the equations at each grid point within a cell.

In the derivations that follow the Fourier transform of the periodic density

is used, i.e.

$$n(\mathbf{r}) = \sum_{\mathbf{G}} \bar{n}(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}\,. \tag{4.27}$$

where $\bar{n}(\mathbf{G})$ is the Fourier transform of the density. All other quantities that are Fourier transformed are spherical functions of the inter-electron separation $|\,\mathbf{r}-\mathbf{r}'\,|$, consequently, use is made of the three-dimensional spherical Fourier transform (FT),

$$\mathrm{FT}[s(r)] = S(\mathbf{G}) = 4\pi \int_0^\infty s(r)\, \frac{\sin(G\,r)}{G}\, r\, dr\,, \tag{4.28}$$

where $s(r)$ is a spherical function, and $G = |\mathbf{G}|$. The generalised pair-correlation function form given by (4.20) is employed in all calculations.

4.3.2 The Sum Rule

The sum rule (4.7) can be re-written using the convolution theorem (4.26) together with the Fourier transform of the density (4.27) as [135, 136],

$$\int n(\mathbf{r}')\, G^{\mathrm{WDA}}[u, \tilde{n}(\mathbf{r})]\, d\mathbf{r}' = \sum_{\mathbf{G}} \bar{n}(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} \times \mathrm{FT}\Big\{G^{\mathrm{WDA}}[u, \tilde{n}]\Big\} = -1\,. \tag{4.29}$$

The Fourier transform of $G^{\mathrm{WDA}}[u; \tilde{n}]$ is obtained using (4.28),

$$\mathrm{FT}\Big\{G^{\mathrm{WDA}}[u; \tilde{n}]\Big\} \;=\; 4\pi \int_0^\infty G^{\mathrm{WDA}}[u; \tilde{n}]\, \frac{\sin(G\,u)}{G}\, u\, du\,, \tag{4.30}$$

substituting $q = \beta(\mathbf{r})G$ and using (4.20) leads to,

$$\mathrm{FT}\Big\{G^{\mathrm{WDA}}[u, \tilde{n}]\Big\} \;=\; \alpha(\mathbf{r})\beta^3(\mathbf{r}) F_1(q)\,, \tag{4.31}$$

where shorthand notation $\alpha(\tilde{n}(\mathbf{r})) = \alpha(\mathbf{r})$ and $\beta(\tilde{n}(\mathbf{r})) = \beta(\mathbf{r})$ is used for brevity. The function $F_1(q)$ is given by,

$$F_1(q) = \frac{4\pi}{q} \int_0^\infty f(u)\, \sin(qu)\, u\, du\,. \tag{4.32}$$

The reciprocal-space representation of the XC sum rule condition therefore becomes,

$$\alpha(\mathbf{r})\beta^3(\mathbf{r}) \sum_{\mathbf{G}} \bar{n}(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} F_1(q) = -1 \tag{4.33}$$

The function $F_1(q)$ is calculated once at the start of a calculation for a sufficient range of q on a uniform grid, and subsequently stored. Particular q-values required during self-consistent cycles are obtained by interpolation.

4.3.3 Energy Density

The reciprocal-space derivation of the WDA energy density $\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ is very similar to that of the sum rule integral given previously. Applying the convolution theorem (4.26) to (4.9) yields,

$$\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r}) = \frac{1}{2}\sum_{\mathbf{G}} n(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} \times \mathrm{FT}\left\{\frac{G^{\rm WDA}[u,\tilde{n}]}{u}\right\} . \tag{4.34}$$

Evaluating the spherical Fourier transform using (4.28) gives,

$$\mathrm{FT}\left\{\frac{G^{\rm WDA}[u,\tilde{n}]}{u}\right\} = \alpha(\mathbf{r})\beta^2(\mathbf{r})\, F_2(q) \tag{4.35}$$

where $F_2(q)$ is defined by,

$$F_2(q) = \frac{4\pi}{q}\int_0^\infty f(u)\sin(qu)\, du\,, \tag{4.36}$$

and is calculated in the same way as $F_1(q)$. The resulting form for $\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ in terms of reciprocal lattice vectors is therefore,

$$\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r}) = \frac{1}{2}\alpha(\mathbf{r})\beta^2(\mathbf{r})\sum_{\mathbf{G}} \bar{n}(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}\, F_2(q)\,. \tag{4.37}$$

4.3.4 The Scalar Fields: $\alpha(\mathbf{r})$ and $\beta(\mathbf{r})$

Before proceeding with the WDA potential, the scalar fields $\alpha(\mathbf{r})$ and $\beta(\mathbf{r})$ that define the WDA hole are first described. As stated previously, these parameters are obtained by demanding that the WDA explicitly satisfies the sum rule and the energy density in the limit of constant density, $n(\mathbf{r}) = n$, i.e.

$$n\int G^{\rm WDA}[u,\tilde{n}(\mathbf{r})]\, d\mathbf{r} = -1\,, \tag{4.38}$$

$$n\int \frac{G^{\rm WDA}[u,\tilde{n}(\mathbf{r})]}{u}\, d\mathbf{r} = \varepsilon_{\rm XC}^{\rm LDA}(n)\,. \tag{4.39}$$

For a homogeneous electron gas, the weighted density is identical to the actual density, i.e. $\tilde{n}(\mathbf{r}) = n = \text{constant}$, therefore relations (4.33) and (4.37) simplify to give,

$$-1 = \tilde{n}\,\alpha(\mathbf{r})\beta^3(\mathbf{r})F_1(0) \tag{4.40}$$

$$\varepsilon_{\text{XC}}^{\text{LDA}}(\tilde{n}) = \frac{1}{2}\tilde{n}\,\alpha(\mathbf{r})\beta^2(\mathbf{r})F_2(0)\,, \tag{4.41}$$

where the constants, $F_1(0)$ and $F_2(0)$ are obtained by setting $q = 0$ in relations (4.32) and (4.36),

$$F_1(0) = 4\pi \int_0^\infty u^2 f(u)\,du \tag{4.42}$$

$$F_2(0) = 4\pi \int_0^\infty u\, f(u)\,du\,. \tag{4.43}$$

Rearranging (4.40) and (4.41) gives rise to the relations that define $\alpha(\mathbf{r})$ and $\beta(\mathbf{r})$:

$$\alpha(\mathbf{r}) = \frac{-1}{\tilde{n}\beta(\mathbf{r})^3\,F_1(0)}\,, \tag{4.44}$$

$$\beta(\mathbf{r}) = \frac{-F_2(0)}{2\,\varepsilon_{\text{XC}}^{\text{LDA}}(\tilde{n})\,F_1(0)}\,. \tag{4.45}$$

4.3.5 WDA Potential

Recall from Sec. 4.2.2 that the WDA potential contains three terms,

$$v_{\text{XC}}^{\text{WDA}}(\mathbf{r}) = v_1(\mathbf{r}) + v_2(\mathbf{r}) + v_3(\mathbf{r})\,, \tag{4.46}$$

where $v_1(\mathbf{r})$ is simply the energy density $\varepsilon_{\text{XC}}^{\text{WDA}}(\mathbf{r})$ given by (4.37). The second and third terms can be calculated by taking their Fourier transforms. Since the Fourier transform of a convolution is just a multiplication of Fourier transforms

$$\text{FT}\Big\{(a * b)(\mathbf{r})\Big\} = A(\mathbf{G})B(\mathbf{G}) \tag{4.47}$$

we are able to write the Fourier transform of $v_2(\mathbf{r})$ as,

$$v_2(\mathbf{G}) = \frac{1}{2}\,\text{FT}\Big\{n(\mathbf{r}')\Big\} \times \text{FT}\left\{\frac{G^{\text{WDA}}[u, \tilde{n}(\mathbf{r}')]}{u}\right\} \tag{4.48}$$

using the expression in (4.35), the Fourier components of $v_2(\mathbf{r})$ are therefore given by,

$$\bar{v}_2(\mathbf{G}) = \frac{2\pi}{\Omega} \int n(\mathbf{r}')\, \alpha(\mathbf{r}')\, \beta^2(\mathbf{r}')\, \mathrm{e}^{-i\mathbf{G}\cdot\mathbf{r}'}\, F_2(q)\, d\mathbf{r}' \tag{4.49}$$

where Ω is the unit cell volume. The real-space value of $v_2(\mathbf{r})$ is therefore obtained by calculating $\bar{v}_2(\mathbf{G})$ at all reciprocal lattice vectors in the cell and then transforming the result back to real space.

The third term, $v_3(\mathbf{r})$, is calculated in exactly the same way, however the reciprocal-space expression is more complicated. Combining relations (4.13) to (4.15), the real-space version can be re-written as,

$$v_3(\mathbf{r}) = -\int n(\mathbf{r}')\, \frac{h_1(\mathbf{r}')}{h_2(\mathbf{r}')}\, G^{\mathrm{WDA}}[\mathbf{r}', \mathbf{r}''; \tilde{n}(\mathbf{r}')]\, d\mathbf{r}'\,, \tag{4.50}$$

where the functions $h_1(\mathbf{r}')$ and $h_2(\mathbf{r}')$ are,

$$h_1(\mathbf{r}') = \int \frac{n(\mathbf{r}')}{|\,\mathbf{r}' - \mathbf{r}''\,|}\, \frac{\partial G^{\mathrm{WDA}}[\mathbf{r}', \mathbf{r}''; \tilde{n}(\mathbf{r}')]}{\partial n(\mathbf{r}')}\, d\mathbf{r}''\,, \tag{4.51}$$

$$h_2(\mathbf{r}') = \int n(\mathbf{r}'')\, \frac{\partial G^{\mathrm{WDA}}[\mathbf{r}', \mathbf{r}''; \tilde{n}(\mathbf{r}')]}{\partial \tilde{n}(\mathbf{r}')}\, d\mathbf{r}''\,. \tag{4.52}$$

As with $v_2(\mathbf{r})$, taking the spherical Fourier transform of $v_3(\mathbf{r})$ leads to,

$$\bar{v}_3(\mathbf{G}) = -\frac{\pi}{\Omega} \int n(\mathbf{r}')\, \alpha(\mathbf{r}')\, \beta^3(\mathbf{r}')\, \frac{h_1(\mathbf{r}')}{h_2(\mathbf{r}')}\, \mathrm{e}^{-i\mathbf{G}\cdot\mathbf{r}'}\, F_1(q)\, d\mathbf{r}'\,. \tag{4.53}$$

The added complication arises from the functions $h_1(\mathbf{r}')$ and $h_2(\mathbf{r}')$. The derivative of $G^{\mathrm{WDA}}[u, \tilde{n}(\mathbf{r})]$ contained in these functions can be written as,

$$\frac{\partial G^{\mathrm{WDA}}[u, \tilde{n}(\mathbf{r})]}{\partial \tilde{n}(\mathbf{r})} = \frac{G^{\mathrm{WDA}}[u, \tilde{n}(\mathbf{r})]}{\alpha(\mathbf{r})}\, \frac{\partial \alpha(\mathbf{r})}{\partial \tilde{n}(\mathbf{r})} + \alpha(\mathbf{r})\, \frac{df(u)}{du}\, \frac{\partial u}{\partial \beta(\mathbf{r})}\, \frac{\partial \beta(\mathbf{r})}{\partial \tilde{n}(\mathbf{r})}\,. \tag{4.54}$$

Substituting this equation into $h_1(\mathbf{r})$ and $h_2(\mathbf{r})$, it is clear that the first term of (4.51) is proportional to $v_1(\mathbf{r})$, while the first term of (4.52) is proportional to the sum rule. The second terms can be calculated using 3D spherical Fourier transforms since they are in the form of convolutions. This leads to,

$$h_1(\mathbf{r}) = \frac{2v_1(\mathbf{r})}{\alpha(\mathbf{r})}\, \frac{\partial \alpha(\mathbf{r})}{\partial \tilde{n}(\mathbf{r})} - \alpha(\mathbf{r})\beta(\mathbf{r})\, \frac{\partial \beta(\mathbf{r})}{\partial \tilde{n}(\mathbf{r})} \sum_{\mathbf{G}} n(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}\, F_3(q)\,, \tag{4.55}$$

$$h_2(\mathbf{r}) = -\frac{1}{\alpha(\mathbf{r})}\, \frac{\partial \alpha(\mathbf{r})}{\partial \tilde{n}(\mathbf{r})} - \alpha(\mathbf{r})\beta^2(\mathbf{r})\, \frac{\partial \beta(\mathbf{r})}{\partial \tilde{n}(\mathbf{r})} \sum_{\mathbf{G}} n(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}}\, F_4(q)\,, \tag{4.56}$$

where

$$F_3(q) = \frac{4\pi}{q}\int_0^\infty u\,\frac{\partial f(u)}{\partial u}\sin(qu)\,du\,, \tag{4.57}$$

$$F_4(q) = \frac{4\pi}{q}\int_0^\infty u^2\,\frac{\partial f(u)}{\partial u}\sin(qu)\,du\,. \tag{4.58}$$

These functions are calculated in the same way as $F_1(q)$ and $F_2(q)$. The derivatives $\partial\alpha(\mathbf{r})/\partial\tilde{n}(\mathbf{r})$ and $\partial\beta(\mathbf{r})/\partial\tilde{n}(\mathbf{r})$ can be obtained from (4.44) and (4.45) as follows,

$$\frac{\partial\alpha(\mathbf{r})}{\partial\tilde{n}(\mathbf{r})} = \frac{1}{\tilde{n}^2\beta(\mathbf{r})^3F_1(0)} + \frac{\partial\alpha(\mathbf{r})}{\partial\beta(\mathbf{r})}\frac{\partial\beta(\mathbf{r})}{\partial\tilde{n}(\mathbf{r})}\,, \tag{4.59}$$

$$\frac{\partial\beta(\mathbf{r})}{\partial\tilde{n}(\mathbf{r})} = \frac{\partial\beta(\mathbf{r})}{\partial\varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})}\frac{d\varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})}{d\tilde{n}(\mathbf{r})}\,, \tag{4.60}$$

substituting $n(\mathbf{r}) = \tilde{n}(\mathbf{r})$ into the LDA potential $v_{\rm XC}^{\rm LDA}(\mathbf{r})$, and rearranging gives,

$$\frac{d\varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})}{d\tilde{n}(\mathbf{r})} = \frac{v_{\rm XC}^{\rm LDA}(\mathbf{r}) - \varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})}{\tilde{n}(\mathbf{r})}\,. \tag{4.61}$$

The required equations are then obtained as,

$$\frac{\partial\alpha(\mathbf{r})}{\partial\tilde{n}(\mathbf{r})} = \frac{1}{\tilde{n}(\mathbf{r})\,\beta(\mathbf{r})^3\,F_1(0)}\left(\frac{1}{\tilde{n}(\mathbf{r})} + \frac{3}{\beta(\mathbf{r})}\frac{\partial\beta(\mathbf{r})}{\partial\tilde{n}(\mathbf{r})}\right)\,, \tag{4.62}$$

$$\frac{\partial\beta(\mathbf{r})}{\partial\tilde{n}(\mathbf{r})} = \frac{F_2(0)}{2\tilde{n}\,F_1(0)}\left(\frac{v_{\rm XC}^{\rm LDA}(\mathbf{r}) - \varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})}{[\varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})]^2}\right)\,. \tag{4.63}$$

Once the Fourier components of $\bar{v}_2(\mathbf{G})$ and $\bar{v}_3(\mathbf{G})$ have been calculated, they are then transformed to real-space via a Fast Fourier Transform (FFT) to give $v_2(\mathbf{r})$ and $v_3(\mathbf{r})$.

A great advantage of the reciprocal space formalism is the ease with which different model pair-correlation functions can be incorporated, since only the function $f(u)$ and its derivative $df(u)/du$ need to be replaced in $F_1(u), F_2(u), F_3(u)$ and $F_4(u)$, and the values of $F_1(0)$ and $F_2(0)$ changed in (4.44) and (4.45).

4.3.6 Implementation Tests

The implementation of the reciprocal-space formulation of the WDA can be checked by performing calculations on a homogeneous electron gas. Firstly, the calculated weighted density at all points should be identically equal to the actual density, i.e. $n(\mathbf{r}) = \tilde{n}(\mathbf{r}) = \tilde{n}$, and the three terms in the WDA potential (4.10) should collapse to the following equations:

$$v_1 = v_2 = \varepsilon_{\rm XC}[\tilde{n}] = \varepsilon_{\rm XC}\,, \tag{4.64}$$

$$v_3 = -\varepsilon_{\rm XC} - \frac{r_s}{3}\frac{\partial \varepsilon_{\rm XC}}{\partial r_s}\,, \tag{4.65}$$

where the $\mathbf{r}$ dependence is dropped due to the spatial invariance of the homogeneous system. The complete WDA potential therefore reduces to the LDA in the homogeneous limit,

$$v_{\rm XC}^{\rm WDA} = \varepsilon_{\rm XC} - \frac{r_s}{3}\frac{\partial \varepsilon_{\rm XC}}{\partial r_s}\,. \tag{4.66}$$

These aspects have been tested and are fulfilled by the implementation. Of course the ultimate test of any functional implementation can be verified by comparing numerical and analytic forces on displaced ions, as performed in Sec. 3.2.3, utilising the Hellmann-Feynman theorem [105, 106]. The implementation of the WDA carried out here indeed fulfils this test.

This test must be performed for every new model pair-correlation function that is implemented. Errors arising from newly implemented models are easy to track down since they can only originate from quantities associated with that model function.

4.4 Algorithm Development

The most computationally intensive part of a WDA calculation is the determination of the weighted density at each point in space. A way of determining the weighted density $\tilde{n}(\mathbf{r})$, is to choose a trial value $\tilde{n}_0(\mathbf{r})$, then calculate the

left-hand side of (4.33) with $\tilde{n}_0(\mathbf{r})$, and iteratively repeat the procedure with different weighted density values until the sum rule is satisfied to within a desired tolerance. This method was implemented initially, however it was found to be very time consuming since the iterative procedure must be performed at all points in space.

In order to make the calculations more practical, an alternative method based on interpolation grids was developed, similar to that described by Singh in Ref. [141]. The new method gives rise to a substantial saving in the computational time, with only a modest increase in computer memory usage. The method is now described in full.

A quantity that occurs repeatedly within the reciprocal-space implementation is the summation over **G**-vectors:

$$\sum_{\mathbf{G}} \bar{n}(\mathbf{G})\, \mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} F_m(q)\,, \tag{4.67}$$

where $m = 1, 2, 3, 4$, corresponds to relations (4.33), (4.37), (4.55) and (4.56) respectively. Instead of re-calculating this expression many times over as in the previous implementation, the summation is computed once for each of the four expressions and stored in the form of look-up tables, using a set of interpolation grids that represent both the spatial coordinate $\mathbf{r}$, and the weighted density, $\tilde{n}$. The set of spatial grid points, $\{\mathbf{r}_i\} = \{x_i, y_j, z_k\}$, are equally separated throughout the three-dimensional cell, representing a mesh of $N_x \times N_y \times N_z$ coordinates. This grid is called the WDA grid to distinguish it from the one used by the CASTEP code to transform between real and reciprocal-space. The dependence on the weighted density is handled using a logarithmic set of values $\{\tilde{n}_i\}$ which contain N_q points. At the beginning of a calculation, the following four-dimensional look-up tables, $T_{n=1,2,3,4}(x_i, y_j, z_k, \tilde{n}_l)$, are created:

$$T_1(x_i, y_j, z_k, \tilde{n}_l) \;=\; \alpha(\tilde{n}_l)\, \beta^3(\tilde{n}_l)\, \sigma(x_i, y_j, z_i)\, F_1[\beta(\tilde{n}_l)\, |\mathbf{G}|]\,, \tag{4.68}$$

$$T_2(x_i, y_j, z_k, \tilde{n}_l) = \frac{1}{2}\alpha(\tilde{n}_l)\,\beta^2(\tilde{n}_l)\,\sigma(x_i, y_j, z_i)\,F_2[\beta(\tilde{n}_l)\,|\mathbf{G}|]\,, \quad (4.69)$$

$$T_3(x_i, y_j, z_k, \tilde{n}_l) = \alpha(\tilde{n}_l)\,\beta(\tilde{n}_l)\frac{\partial\beta\tilde{n}_l}{\partial\tilde{n}_l}\,\sigma(x_i, y_j, z_i)\,F_3[\beta(\tilde{n}_l)\,|\mathbf{G}|]\,, \quad (4.70)$$

$$T_4(x_i, y_j, z_k, \tilde{n}_l) = \alpha(\tilde{n}_l)\,\beta^2(\tilde{n}_l)\frac{\partial\beta\tilde{n}_l}{\partial\tilde{n}_l}\,\sigma(x_i, y_j, z_i)\,F_4[\beta(\tilde{n}_l)\,|\mathbf{G}|]\,, \quad (4.71)$$

where the dependence on the spatial coordinates enters into $\sigma(x_i, y_j, z_i)$, which is given by,

$$\sigma(x_i, y_j, z_i) = \sum_{\mathbf{G}} n(\mathbf{G})\,\mathrm{e}^{i(G_x x_i + G_y y_j + G_z z_k)}\,. \quad (4.72)$$

Relations T_1 and T_2 represent the sum rule and energy density, while T_3 and T_4 are used to determine the right-hand side of $h_1(\mathbf{r})$ and $h_2(\mathbf{r})$ respectively. So, each look-up table, consisting of $N_x N_y N_z N_{\tilde{n}}$ values, need only be calculated once at the beginning of a self-consistent calculation. In the case of a geometry optimisation calculation, the tables must be re-calculated whenever the set of **G**-vectors changes.

Since a given position **r** on the CASTEP grid will not correspond to a point on the WDA grid, six-point Lagrange interpolations are performed using values from the look-up table in order to calculate the required quantity at the actual position **r**. An interpolation must obviously be carried-out in each of the three spatial dimensions. To determine the precise value of the weighted density at a point **r**, the spatial interpolations are performed using values from T_1 that surround the point **r**, a set of weighted density grid values are then cycled through until the actual weighted density $\tilde{n}$ (i.e. the one that yields the sum rule) is bound by two values in the grid, $\tilde{n}_{l-1}(\mathbf{r})$ and $\tilde{n}_l(\mathbf{r})$. An inverse interpolation is then performed to obtain the precise value of the weighted density that yields $T_1 = -1$. The subsequent determination of $\varepsilon_{\mathrm{XC}}^{\mathrm{WDA}}(\mathbf{r})$ and $v_{\mathrm{XC}}^{\mathrm{WDA}}(\mathbf{r})$ proceeds in much the same way but is simplified by the fact that $\tilde{n}(\mathbf{r})$ is already known, so the four-dimensional interpolations on T_2, T_3 and T_4 can be performed directly.

The advantage of using this method is that the WDA spatial grid is smaller than the CASTEP grid, which reduces the number of time-consuming reciprocal-space summations (4.67). However, the interpolation grids should be sufficiently fine enough so that the inhomogeneities in the energy density and potential are accurately described. To establish this, convergence tests on the total energy of the system with respect to N_x, N_y, N_z and $N_{\tilde{n}}$ must be performed.

4.5 Self-Consistent WDA Calculations

As a first application of this WDA implementation, the lattice constant and bulk modulus of C-diamond, Si, Ge, GaAs and Al are determined using the WDA. The emphasis here is to provide a demonstration of how to perform self-consistent calculations using this implementation of the WDA. For all of the WDA calculations, the following model pair-correlation function is used,

$$G^{\mathrm{WDA}}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})] = \alpha(\tilde{n})\,\mathrm{e}^{-[|\mathbf{r}-\mathbf{r}'|/\beta(\tilde{n})]^2}\,, \tag{4.73}$$

which according to relation (4.20) corresponds to the function $f(u)$ with the form,

$$f(u) = \mathrm{e}^{-u^2}\,. \tag{4.74}$$

This happens to be the simplest physical model for a pair-correlation function, and will be discussed in detail in subsequent chapters.

4.5.1 Interpolation Grid Convergence

As with all plane-wave calculations the usual convergence tests for the kinetic energy cutoff and number of k-points must be performed. Experience shows that for a given system the specific convergence values will be the same for all functionals, including the WDA, so it is easiest to use the LDA for these tests.

The first quantity to establish in a WDA calculation is the maximum weighted density interpolation value $\tilde{n}_{max}$ required for the system under study. There is no hard and fast rule to predict this value for a given system, although the greater the inhomogeneity in the density, and the larger the number of electrons in the unit cell, will generally mean that a higher value of $\tilde{n}_{max}$ will be needed. Once $\tilde{n}_{max}$ is established, the second stage is to perform a convergence test on the number of weighted density grid points $N_{\tilde{n}}$ - single-point energy calculations are performed, and the results are plotted against increasing $N_{\tilde{n}}$. The converged value is determined in the usual fashion, i.e. when the curve becomes flat. Fig. 4.1 shows the convergence of $N_{\tilde{n}}$ for the solids chosen here.

The third and final stage it to converge the three-dimensional spatial grid to obtain N_x, N_y and N_z. For bulk solids this can be performed for just one of the unit cell directions and the same converged value used for the other two directions. Obviously for systems that are distinctly different in particular directions, such as a surface calculation where the bulk directions are very different from that perpendicular to the surface, then convergence tests will be required in the other directions. Fig. 4.2 shows the convergence test of N_x for the primitive cells of C-diamond, Si, and GaAs, and the unit cell of Al. It can be seen from this figure that the systems with more rapidly varying densities such as GaAs require more grid points than other systems. Note that the convergence for Al was performed on the 4 atom non-primitive unit cell, and so actually requires the least number of spatial grid points per unit volume. This is to be expected since the metallic density of Al varies most slowly.

Table. 4.1 gives the chosen values of $\tilde{n}_{max}$, together with the converged values for the weighted density and spatial interpolation grid sizes for each of the five systems, which were obtained from the convergence tests presented in Figs. 4.1 and 4.2. The calculations were performed using LDA ultrasoft pseudopotentials.

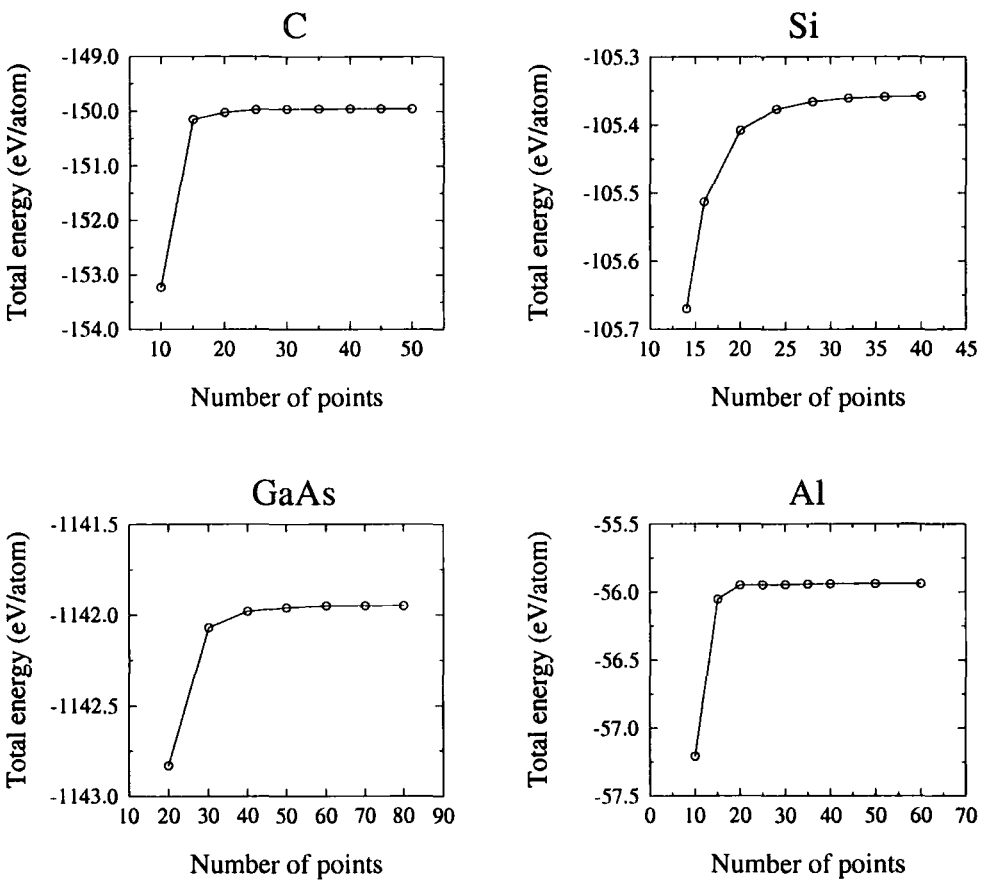

Figure 4.1: Convergence of the total energy, $E[n(\mathbf{r})]$, calculated using the WDA, with respect to the number of WDA weighted density interpolation points $N_{\tilde{n}}$, for bulk C-diamond, Si, GaAs and Al.

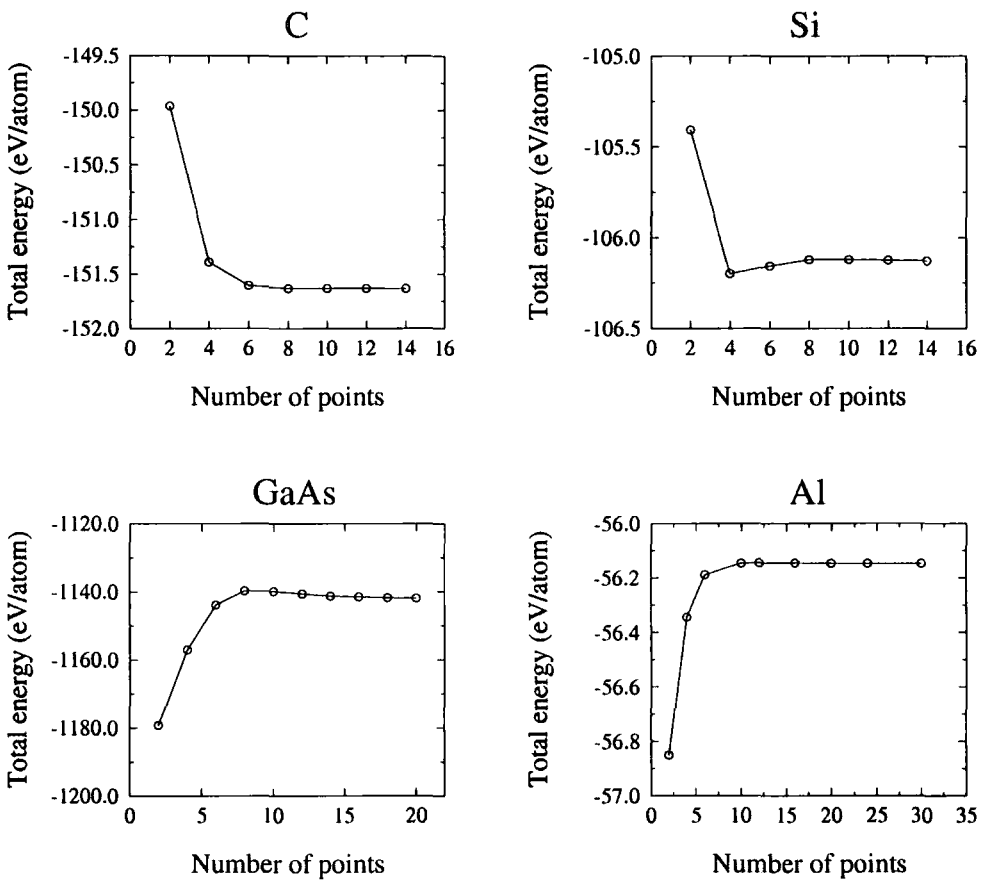

Figure 4.2: Convergence of the total energy, $E[n(\mathbf{r})]$, calculated using the WDA, with respect to the number of WDA spatial interpolation grid points N_i along one of the crystal directions, for C, Si, GaAs and Al. A non-primitive unit cell is used for Al.

Table 4.1: Values of the WDA interpolation grids used in the self-consistent calculations determined from the convergence tests presented in Figs. 4.1 and 4.2. The value of $\tilde{n}_{max}$ is also given for each system.

	$N_x \times N_y \times N_z$	$N_{\tilde{n}}$	$\tilde{n}_{max}$
C	$8 \times 8 \times 8$	25	100
Si	$10 \times 10 \times 10$	30	100
Ge	$10 \times 10 \times 10$	40	600
GaAs	$16 \times 16 \times 16$	50	2000
Al	$16 \times 16 \times 16$	25	100

4.5.2 Results

The self-consistent determination of the lattice constants and bulk-moduli were computed using the converged grid values presented in Table 4.1. LDA ultrasoft pseudopotentials are also employed in the WDA calculations.

The WDA implementation becomes exact when the spatial grid is the same size as the CASTEP FFT grid. For all of the semiconductor calculations the FFT grid size was $24 \times 24 \times 24$, using the converged cutoff energy of 300 eV. Comparing this size with the considerably smaller, although converged, WDA grids in Table 4.1 demonstrates that the interpolation scheme offers a substantial computational saving over the formally exact calculations.

As a confirmation of the spatial grid convergence tests, the lattice constant and bulk modulus of Si has been calculated for five different sizes with increasing precision. The results, presented in Table 4.2, demonstrate that the lattice constant and bulk modulus are sufficiently converged with a grid of size, $N_x = N_y = N_z = 10$, which is in agreement with the original total energy convergence test.

Fig. 4.3 shows the results of the Murnaghan interpolations for each of the solids calculated using the WDA, and Table 4.3 shows the lattice constants and bulk moduli derived from these fits. Corresponding LDA and PBE-GGA

Table 4.2: The lattice constant (a_0) and bulk modulus (B_0) of Si calculated using various spatial grid sizes.

	$N_x = N_y = N_z$					
	2	4	6	8	10	12
a_0 (Å)	5.164	5.414	5.433	5.436	5.437	5.438
B_0 (GPa)	153.4	96.4	88.9	88.7	87.9	87.5

values are also given, which were determined using LDA pseudopotentials to be consistent with the WDA calculations. The WDA shows promising results in general, nearly all lattice constants are in better agreement with experiment than both the LDA and the GGA. The exception is for GaAs where the poorer result is probably related to the use of the LDA pseudopotential. To give an indication of the computational times involved for these semiconductor systems, on a Compaq Alpha XP1000 667MHz computer, a single self-consistent calculation with the LDA or the GGA takes between $2-6$ minutes, whereas the WDA takes between $30-80$ minutes depending on the size of the unit cell, using the grid sizes presented in Table 4.1.

Table 4.3: Lattice constants and bulk moduli of several solids determined using the LDA, PBE-GGA and the WDA. LDA ultrasoft pseudopotentials are used for all three functionals.

	Lattice constant (Å)				Bulk modulus (GPa)			
	LDA	PBE	WDA	Expt.	LDA	PBE	WDA	Expt.
C	3.53	3.53	3.56	3.57	457	447	424	442
Si	5.38	5.38	5.44	5.43	97	93	88	98.8
Ge	5.54	5.55	5.60	5.66	78	75	69	76.8
GaAs	5.57	5.65	5.53	5.65	75	65	88	74.8
Al	3.98	3.98	3.98	4.05	82	83	83	77.3

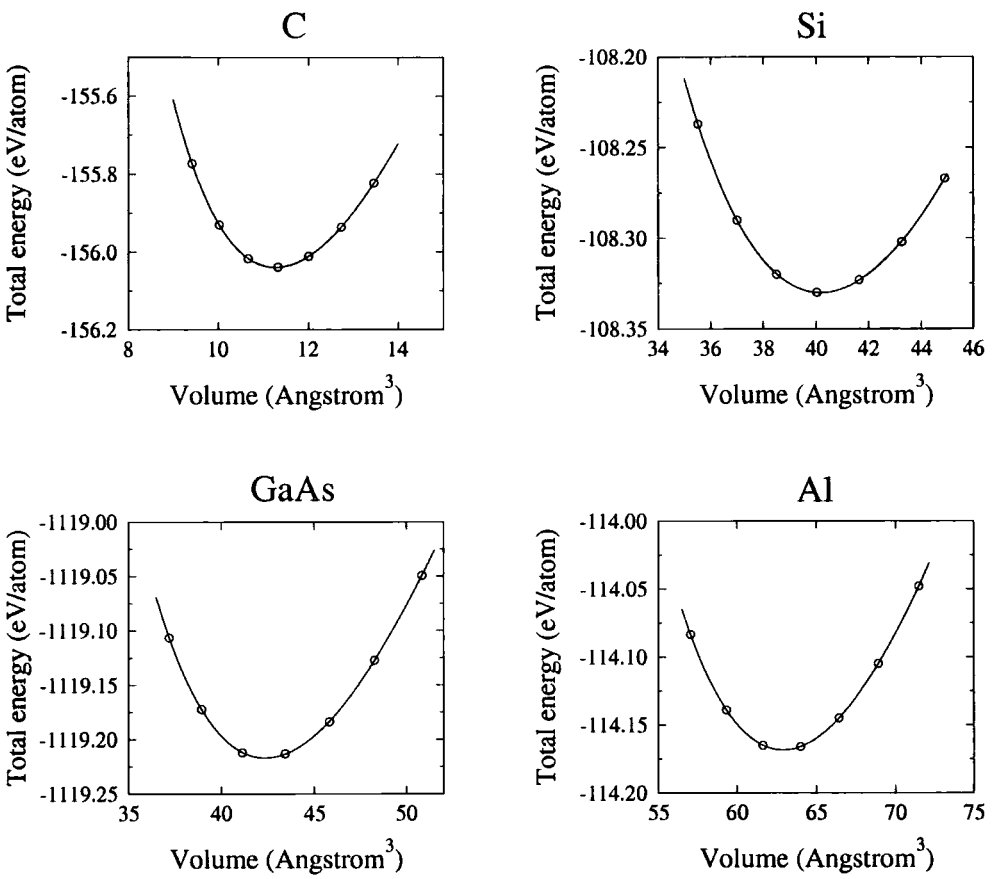

Figure 4.3: Single-point energy calculations performed using the WDA (open circles) and the corresponding Murnaghan fits (solid lines). The calculations were performed using LDA pseudopotentials and the WDA grid sizes shown in Table 4.1.

The importance of the pseudopotential consistency issue when testing functionals is re-iterated in the PBE results since for most systems PBE yields very similar properties as the LDA. However this is certainly not the case when consistent pseudopotentials are employed. In this case the usual correction to the LDA overbinding is observed from PBE and the structures are almost identical to those found for the PW91 functional as given in Tables 3.1 and 3.2 in Chapter 3. The WDA structures similarly undergo a lattice expansion when PBE pseudopotentials are employed.

4.6 Summary

In summary, the WDA is most efficiently realised within a reciprocal-space formalism because of the exploitation of the convolution theorem, consequently the CASTEP plane-wave pseudopotential code was used in the implementation of the WDA performed here. However, even within this framework the WDA is still prohibitively slow for practical use. To alleviate this problem a faster algorithm was devised that is based around interpolation grids which are converged in a similar manner as for the kinetic energy cutoff and k-point sampling - before the main calculations are conducted. This allows a controlled compromise between accuracy and computational resources, since the WDA equations are exact when the spatial grid is the same size as the CASTEP FFT grid, with a sufficiently fine weighted density grid. Faster results can often be obtained at the expense of a negligible reduction in accuracy when smaller WDA grids are used.

As a test of the implementation, the interpolation grid sizes of some simple materials was established through convergence tests and their corresponding structural properties were calculated using the WDA. Although this was primarily a demonstration of the convergence tests, an initial assessment of the WDA shows that it is very promising in comparison to the LDA and a popular GGA, taking account of the inconsistent pseudpotentials.

Now that an effective computational scheme has been devised for the WDA, it is possible to explore and develop the WDA form. An important area to investigate is the effect of using different analytic models for $G^{\mathrm{WDA}}[\mathbf{r}, \mathbf{r}'; \tilde{n}]$. This is the subject of chapter 5.

Chapter 5

Investigation of Model Pair-Correlation Functions

5.1 Introduction

The WDA only makes an approximation to the form of the coupling-constant averaged pair-correlation function, $g_{\rm XC}(\mathbf{r},\mathbf{r}')$, or in common WDA notation the function, $G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$, and so is the single most important quantity to investigate in the theory of the WDA. Despite such significance, little effort has been invested in a systematic investigation of approximate model functions, despite there being a multitude of physically plausible forms that can be constructed from the rather limited number of exact known constraints. Of the work that has been performed in this area of the WDA, conflicting views have emerged regarding the importance of the precise shape of the model pair-correlation function. This is because the known exact constraints only concern the behaviour of the pair-correlation function at the limits of zero and infinite inter-electron separation, which leaves great scope for the form of $G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$ in the range between these limits. Fritsche [142] suggests that the sum rule on the XC hole together with the condition that $g_{\rm XC}(\mathbf{r},\mathbf{r}') \to 1$, as $\mathbf{r}' \to \infty$, exert such strong constraints that differences

in the model form will have relatively little influence on physical quantities such as the total energy. The majority of other research indicates otherwise: Singh [141] obtained different values for lattice constants and bulk moduli of several solids using the ansatz of Gunnarsson and Jones [138] $G^{\mathrm{GJ}}[\mathbf{r},\mathbf{r};\tilde{n}(\mathbf{r})]$, given by (4.19), and the homogeneous electron gas pair-correlation function $g_{\mathrm{XC}}^{\mathrm{hom}}(\mathbf{r},\mathbf{r}')$, where the parametrisation of Perdew and Wang [143] was used for $g_{\mathrm{C}}^{\mathrm{hom}}(\mathbf{r},\mathbf{r}')$. Krüger *et al.* [144] calculated the bandstructure of Si using $G^{\mathrm{GJ}}[\mathbf{r},\mathbf{r};\tilde{n}(\mathbf{r})]$ and $g_{\mathrm{X}}^{\mathrm{hom}}(\mathbf{r},\mathbf{r}')$ given in (4.17), and obtained improved band energies with $G^{\mathrm{GJ}}[\mathbf{r},\mathbf{r};\tilde{n}(\mathbf{r})]$. They concluded that the detailed shape of the model pair-correlation function does lead to important quantitative differences.

This chapter investigates a range of model pair-correlation functions with intent to clarify this issue. The models will be analysed at the macroscopic scale by calculating bulk properties of Si, and also at the microscopic level by examining XC holes for electrons at various points within the Si crystal, and comparing the results with data from variational Monte Carlo simulations [145].

5.2 The Model Functions

Approximate models for the WDA pair-correlation function, other than the one proposed by Gunnarsson and Jones and the homogeneous electron gas result, have been employed previously. Examples include the single parameter Gaussian function for exchange adopted by Sadd and Teter [146, 147],

$$g_{\mathrm{X}}(\mathbf{r},\mathbf{r}') = 1 - \mathrm{e}^{-\alpha(\mathbf{r})|\,\mathbf{r}-\mathbf{r}'\,|^2}, \tag{5.1}$$

where $\alpha(\mathbf{r})$ is determined from the sum rule. Similarly Gritsenko *et al.* [148] tested a similar function for exchange as (5.1) but with a value of 3/2 in the exponent. A more detailed study was performed by Pedroza [149], who

proposed two model functions, $g^1_{\rm XC}(\mathbf{r},\mathbf{r}')$ and $g^2_{\rm XC}(\mathbf{r},\mathbf{r}')$, with the form,

$$g^1_{\rm XC}(\mathbf{r},\mathbf{r}') = 1-[a(\mathbf{r})+b(\mathbf{r})|\,\mathbf{r}-\mathbf{r}'\,|]e^{-c(\mathbf{r})|\,\mathbf{r}-\mathbf{r}'\,|^2}, \tag{5.2}$$

and

$$g^2_{\rm XC}(\mathbf{r},\mathbf{r}') = 1-[a(\mathbf{r})+b(\mathbf{r})|\,\mathbf{r}-\mathbf{r}'\,|]e^{-c(\mathbf{r})|\,\mathbf{r}-\mathbf{r}'\,|}, \tag{5.3}$$

where $a(\mathbf{r})$ is obtained from an arbitrarily defined function, $\zeta/(1+\gamma r_s)$, with constants ζ and γ. The other parameters, $b(\mathbf{r})$ and $c(\mathbf{r})$, are obtained by satisfying the sum rule and energy density for the homogeneous electron gas. On application to some light atoms, Pedroza found that both functions gave similar total XC energies for single shell atoms, whereas for multishell atoms the energies were more sensitive to the particular model. The differences were attributed to the description of XC effects in the intershell regions by the WDA. Both functions, however, gave good improvement over the LDA for the total XC energy of the atoms considered.

One of the most comprehensive studies of WDA pair-correlation functions was performed by Charlesworth [150], who calculated the lattice constant and bulk modulus of several bulk solids using twelve simple model functions of the two parameter form suggested by Gunnarsson and Jones [138]. Charlesworth reported extremely poor results for every WDA function, which in most cases were considerably worse than the LDA. He concluded that the WDA is fundamentally flawed because of the spherical approximation contained in the model pair-correlation functions, i.e. using the quantity $|\mathbf{r}-\mathbf{r}'|$ instead of separate dependence on $\mathbf{r}$ and $\mathbf{r}'$. However, judging from the self-consistent results obtained at the end of Chapter 4, this conclusion seems questionable. For example, using LDA pseudopotentials and the same model for $G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$, Charlesworth calculated the lattice constant of Si and Ge to be 5.63 and 6.12Å respectively, whereas the corresponding results presented in Table 4.3 yield 5.43 and 5.60Å respectively. The wide variation in these results means that a further investigation is required. Consequently, the same twelve model functions used by Charlesworth are now re-examined.

The twelve functions consist of four distinct mathematical groups and have the general form $G_i^{\mathrm{WDA}}(u,\alpha) = \alpha f_i(u)$, as given by (4.20), so are characterised by the analytic expression $f_i(u)$, which are displayed in Table 5.1. Groups 1 and 2 have short ranged exponential decaying forms, whereas the functions in Groups 3 and 4 exhibit a power law decay and so possess long-range asymptotic tails. The form originally proposed by Gunnarsson and Jones constitutes the fourth group, and the specific $G^{\mathrm{GJ}}(u,\alpha)$ function is given by $f_{11}(u)$.

Table 5.1: The twelve model pair-correlation functions employed in the WDA calculations. They have the general form $G^{\mathrm{WDA}}(u,\alpha) = \alpha f(u)$, so the characteristic expressions $f(u)$ are given.

$f_i(u)$	Function	$f_i(u)$	Function
Group1		Group3	
$f_1(u)$	e^{-u^2}	$f_7(u)$	$1/(1+u^4)$
$f_2(u)$	$(1+u^2)\,\mathrm{e}^{-u^2}$	$f_8(u)$	$1/(1+u^5)$
$f_3(u)$	$(1+u^2+\frac{1}{2}u^4)\,\mathrm{e}^{-u^2}$	$f_9(u)$	$1/(1+u^6)$
Group2		Group4	
$f_4(u)$	e^{-u^4}	$f_{10}(u)$	$(1-\mathrm{e}^{-1/u^4})$
$f_5(u)$	$(1+u^4)\,\mathrm{e}^{-u^4}$	$f_{11}(u)$	$(1-\mathrm{e}^{-1/u^5})$
$f_6(u)$	$(1+u^4+u^8)\,\mathrm{e}^{-u^4}$	$f_{12}(u)$	$(1-\mathrm{e}^{-1/u^6})$

The variety of shapes encompassed by this set of models is demonstrated in Fig. 5.1, which plots $g_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = g_{\mathrm{xc}}(r)$ as a function of the separation, $r = |\mathbf{r}-\mathbf{r}'|$, for a homogeneous electron gas with average density corresponding to $r_s = 1.0a_0$. Each of the four groups have a distinctive shape, and within each group the functions exhibit a variety of spatial ranges, so the effect of shape and spatial extent on physical properties can be systematically analysed. The shape of the Gaussian functions in Fig. 5.1(a) are most like the homogeneous gas function function that gives rise to the LDA.

The fourth order Gaussians shown in Fig. 5.1(b) display very similar spatial extents, which is reflected by the small range of on-top values $g_{\rm xc}(0)$, consequently the effect of small changes in the pair-correlation function can also be examined. These functions however have unphysical on-top values at high density since they are greater than 0.5, which violates the upper bound of condition (2.26). This deficiency also occurs for the functions f_3 and f_{12}. In contrast to this, function f_7 of group 3 (the Lorentzian functions) and f_{10} of the Gunnarsson-Jones type models display unusually low on-top values, with f_7 actually becoming negative as $r \to 0$, which violates the lower bound in (2.26).

Another inadequacy of these functions is that they possess a zero first derivative when evaluated at $r' = r$, and so do not satisfy the cusp condition attributed to Kimball [151], that arises from correlation interactions,

$$\left.\frac{dg_{\rm xc}(\mathbf{r},\mathbf{r}')}{dr}\right|_{\mathbf{r}\to\mathbf{r}'} = \left. g_{\rm xc}(\mathbf{r},\mathbf{r}')\right|_{\mathbf{r}\to\mathbf{r}'}. \tag{5.4}$$

where $g_{\rm xc}(\mathbf{r},\mathbf{r}')$ in this case is the exact (unknown) pair-correlation function. This condition becomes more important in low density environments due to the predominance of correlation interactions over exchange, and also in spin-unpolarised systems [152]. Nevertheless, the wide range of functions considered will provide a useful starting point for investigating the effects of systematic changes in $G^{\rm WDA}[\mathbf{r},\mathbf{r}';\tilde{n}(\mathbf{r})]$ on calculated properties, and obviously in testing Charlesworths claims about the WDA.

5.3 Bulk Properties of Si

All self-consistent WDA calculations employ an ultrasoft pseudopotential for Si generated using the LDA. Although this is not ideal for the reasons stated in Sec. 3.3.6, the fact that it is an LDA pseudopotential makes it consistent with the work of Charlesworth. The convergence of the spatial and weighted density grids in the WDA is the same for all pair-correlation

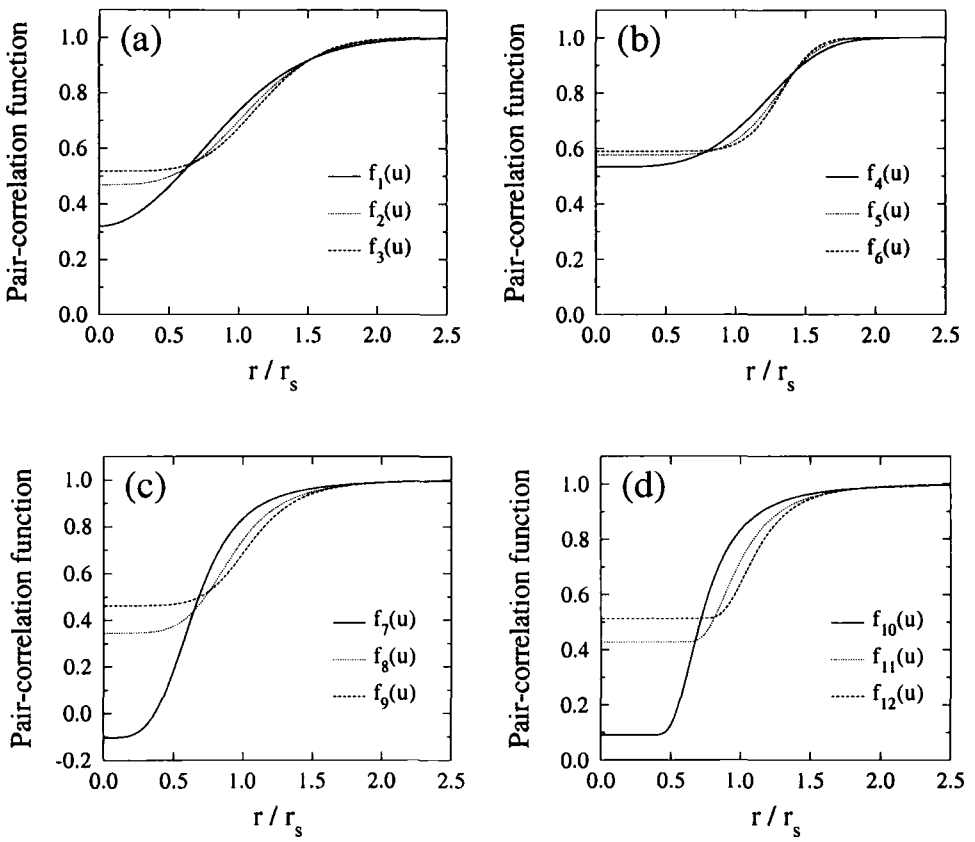

Figure 5.1: Pair-correlation functions $g_{\mathrm{XC}}(r)$ calculated for a homogeneous electron gas with $r_s = 1$ using WDA functions (a) f_1,f_2 and f_3, (b) f_4,f_5, and f_6, (c) f_7,f_8, and f_9, and (d) f_{10},f_{11} and f_{12}. Note the slight change of scale in (c) due to the function $f_7(u)$ going negative as $r \to 0$.

functions, consequently the same grid sizes are used in all calculations. As an example, Fig. 5.2 shows the convergence of the total energy with respect to the WDA spatial grid size $\{\mathbf{r}_i\}$, using the model functions $f_1(u)$ and $f_6(u)$, it is clear that the convergence is almost identical for the two models. A total of 1000 points were used for the spatial grid, i.e $N_x = N_y = N_z = 10$. For the weighted density grid, 30 points were used on the logarithmic scale with a maximum value $\tilde{n}_{max} = 100$. These values converge lattice constants to less than 0.01 Å and total energies to 1 meV/atom or less.

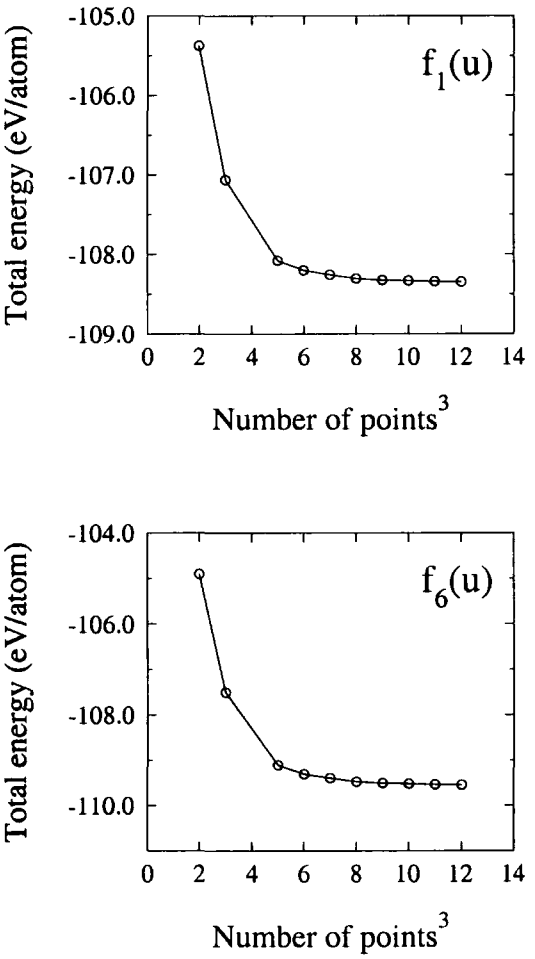

Figure 5.2: Total energy convergence of the WDA spatial grid $\{r_i\}$ (total number of grid points) for the primitive Si cell, using the functions $f_1(u)$ and $f_6(u)$.

5.3.1 Structural Properties

Table 5.2 displays the lattice constant and bulk modulus of Si calculated using the twelve model pair-correlation functions in the WDA. It is clear that there is a wide range of values indicating that the WDA is definitely sensitive to the choice of model function. The WDA is most sensitive to the functions with long ranged asymptotic tails as these give the largest spread of results. The lattice constant and bulk modulus given by the group 4 functions varies by 0.13 Å and 22 GPa respectively, compared with variations of 0.01 Å and 2 GPa for the functions in group 2. The degree of variation in the results within each group closely resembles the qualitative differences in the pair-correlation functions displayed in Fig. 5.1 for the homogeneous electron gas.

Overall, the simple Gaussian functions of group 1 perform best since they give lattice constants in closest agreement with experiment and also very reasonable bulk moduli. Charlesworth [150] found all WDA forms to produce fairly significant overestimates of the lattice constants, and consequently large underestimates of the bulk modulus in relation to experiment, with mean errors of +0.24 Å and −33 GPa for the lattice constant and bulk modulus respectively. Charlesworth's results are also given in Table 5.2. The results obtained here are in contrast. Firstly five of the functions actually yield lattice constants that are shorter than experiment, and the mean errors determined here are +0.02 Å and −12 GPa - significantly smaller than the values quoted by Charlesworth. There is no evidence of an inherent deficiency in the WDA method, at least for the structural quantities examined. It should also be noted that Charlesworth obtained uncharacteristic results with the LDA for all of the solids studied. Lattice constants for example were nearly always larger than experiment. This is at variance with common observations of the LDA.

Despite such quantitative differences, the results obtained here share the

Table 5.2: The equilibrium lattice constant (in Å) and bulk modulus (in GPa) of Si, calculated using the twelve WDA model pair-correlation functions. The columns headed "Ref. [150]" correspond to the values obtained by Charlesworth. The bracketed values are percentage differences from experiment. LDA and experimental values are also given for comparison.

	Lattice constant (Å)		Bulk modulus (GPa)	
$f_i(u)$	This work	Ref. [150]	This work	Ref. [150]
$f_1(u)$	5.44 (+0.2%)	5.63 (+3.7%)	87.9 (−11%)	66.9 (−32%)
$f_2(u)$	5.42 (−0.2%)	5.57 (+2.6%)	91.6 (− 7%)	74.9 (−24%)
$f_3(u)$	5.41 (−0.4%)	5.54 (+2.0%)	93.4 (− 6%)	78.9 (−20%)
$f_4(u)$	5.40 (−0.5%)	5.53 (+1.8%)	94.9 (− 4%)	81.4 (−18%)
$f_5(u)$	5.39 (−0.7%)	5.50 (+1.3%)	96.5 (− 2%)	85.4 (−14%)
$f_6(u)$	5.39 (−0.7%)	5.49 (+1.1%)	97.1 (− 2%)	86.9 (−12%)
$f_7(u)$	5.57 (+2.6%)	6.15 (+13%)	70.6 (−29%)	21.4 (−78%)
$f_8(u)$	5.47 (+0.8%)	5.72 (+5.3%)	81.7 (−17%)	56.8 (−43%)
$f_9(u)$	5.44 (+0.2%)	5.61 (+3.3%)	87.9 (−11%)	69.4 (−30%)
$f_{10}(u)$	5.56 (+2.4%)	6.04 (+11%)	67.5 (−32%)	28.5 (−71%)
$f_{11}(u)$	5.46 (+0.6%)	5.67 (+4.4%)	84.0 (−15%)	62.0 (−37%)
$f_{12}(u)$	5.43 (+0.0%)	5.58 (+2.8%)	89.9 (− 9%)	73.6 (−26%)
LDA	5.38 (−1.0)	5.49 (+1.1%)	97.1 (−2)	88.9 (−10%)
Expt.	5.43		98.8	

same qualitative trends as those reported by Charlesworth: on moving from the shortest to the longest ranged function in each group, the bonding becomes stronger, causing lattice constants to decrease and bulk moduli to increase. In agreement with Charlesworth, we find that the worst results come from the functions, f_7 and f_{10}, where the bonding is found to be considerably weaker than the other models.

5.3.2 Electronic Structure

Table 5.3 shows the minimum band gap of Si calculated with each model pair-correlation function. Again, a significant range of results is obtained. All of the pair-correlation functions yield gaps that are smaller than experiment, sometimes even smaller than the LDA, except the functions f_7 and f_{10} which produce distinctly large gaps that actually overestimate experiment. Within each group the size of the gap decreases as the spatial extent of each pair-correlation function increases. All trends remain the same when calculated at the experimental lattice constant, as observed from the bracketed values in Table 5.3. Charlesworth did not calculate bandstructures and so comparisons cannot be made.

The bandstructure obtained from WDA - f_7 and the LDA, are superimposed in Fig. 5.3 for comparison. This particular pair-correlation function gives rise to a significant shift in most of the conduction bands across the entire Brillouin zone with respect to the LDA. The same effect is observed for WDA - f_{10}, although the magnitude is slightly smaller. A clear illustration of the effect of the pair-correlation functions on the electronic structure is given by Fig. 5.4, which shows the energy gaps between the lowest conduction band states at the L, Γ and X points, and the top of the valence band for the LDA, WDA and experiment. For the WDA, only the results for the first pair-correlation function in each group is presented, since within each group a simple trend is exhibited whereby the gap size decreases for the second and

Table 5.3: The minimum electronic band gap of Si (in eV) yielded by the WDA using the twelve pair-correlation functions, calculated at the equilibrium lattice constant, and the experimental lattice constant (in brackets).

Function	E_g
$f_1(u)$	0.55 (0.54)
$f_2(u)$	0.41 (0.42)
$f_3(u)$	0.34 (0.36)
$f_7(u)$	1.32 (1.20)
$f_8(u)$	0.71 (0.69)
$f_9(u)$	0.51 (0.50)
$f_4(u)$	0.30 (0.33)
$f_5(u)$	0.23 (0.27)
$f_6(u)$	0.21 (0.24)
$f_{10}(u)$	1.21 (1.11)
$f_{11}(u)$	0.62 (0.61)
$f_{12}(u)$	0.44 (0.45)
LDA	0.44 (0.49)
Expt.	1.17

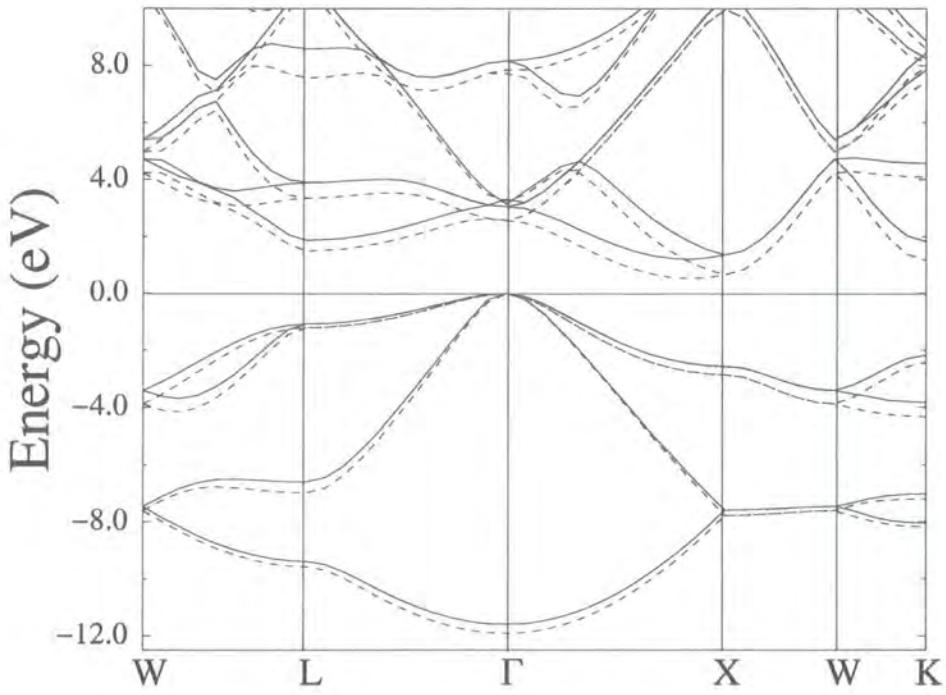

Figure 5.3: The bandstructure of Si calculated at the experimental lattice constant using the WDA with model function f_7 (solid line) and the LDA (dotted line). The top of the valence bands have been lined up at the Γ point.

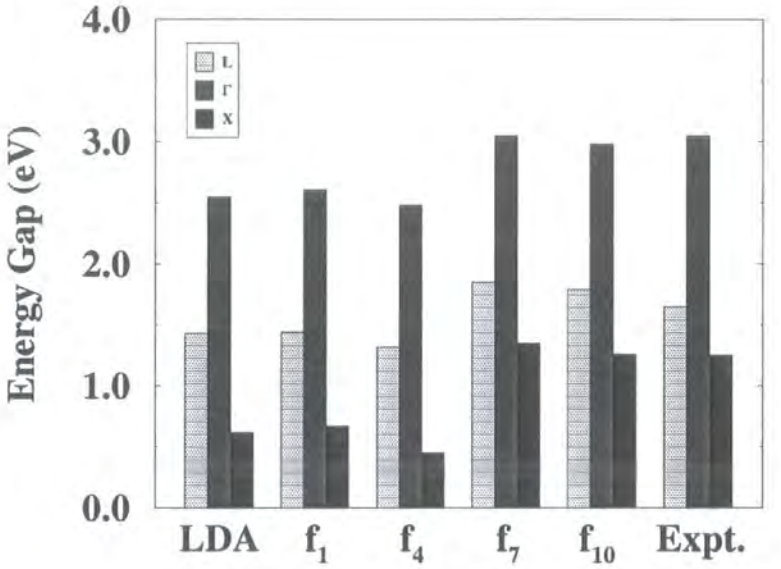

Figure 5.4: Energy gaps for Si (in eV) between the lowest conduction band states at the L, Γ and X points in the Brillouin zone, and the top of the valence band, calculated with the LDA, and the WDA with the functions f_1, f_4, f_7 and f_{10}. The experimental values are taken from Refs. [153] and [109].

third functions. It is evident from this figure that the WDA functions f_7 and f_{10}, lead to significant increases in all three band gaps over the other WDA pair-correlation functions and also the LDA, and are in very good agreement with experiment.

The energy eigenvalues obtained from the LDA, WDA - f_{10}, and experiment, at high-symmetry points for Si at the experimental lattice constant are listed in Table 5.4 for valence and conduction bands. Again, the WDA - f_{10} method yields results in good agreement with experiment, however the value of the Γ_{1v} eigenvalue, which corresponds to the valence band-width, is disappointingly smaller than the LDA result, and consequently even smaller than experiment. This is a general feature of the WDA and so can be explained by an increase and subsequent overestimation of the localisation in the valence

Table 5.4: Band energies of Si (in eV) at the experimental lattice constant calculated with the LDA and the WDA using function f_{10}, and compared with experimental values obtained from Refs. [109, 155].

	LDA	WDA - f_{10}	Expt.
Γ_{1v}	-11.91	-11.67	-12.5 ± 0.6
Γ_{25v}	0.00	0.00	0.00
Γ_{15c}	2.55	2.97	3.40
Γ_{2c}	3.21	3.40	4.15 ± 0.05
X_{1v}	-7.79	-7.64	–
X_{4v}	-2.82	-2.60	-2.9
X_{1c}	0.62	1.26	1.3
X_{4c}	10.01	10.28	–
L_{2v}	-9.60	-9.44	-9.3 ± 0.4
L_{1v}	-6.96	-6.64	-6.7 ± 0.2
L_{3v}	-1.18	-1.09	-1.2 ± 0.2
L_{1c}	1.43	1.79	2.04 ± 0.06
L_{3c}	3.31	3.81	3.9

electron states, caused by the smaller self-interaction effects inherent in the WDA. The same qualitative results have been observed in the bandstructures determined with the exact Kohn-Sham exchange method [154].

Bandstructure calculations of other materials conducted with the WDA using functions f_7 and f_{10}, similarly give rise to upward shifts of the conduction bands throughout the Brillouin zone, bringing band gaps in closer agreement with experiment. Examples of band gaps for other materials are presented in Table 5.5 for the f_{10} pair-correlation function. There are two notable exceptions were the gap is not improved with the WDA - Ge and GaAs. However this is actually expected since it has been shown that the gaps of these systems originate from the core electron interactions [154, 156]

Table 5.5: Minimum band gaps (in eV) of several materials calculated with WDA - $f10$ at the experimental lattice constants.

	LDA	WDA - f_{10}	Expt.
C	4.12	4.91	5.48
Ge	0.03	0.00	0.71
GaAs	0.32	0.22	1.52
GaN	1.74	2.29	3.45
LiCl	6.18	6.95	9.40
$KTaO_3$	1.63	2.60	3.80

which are not properly accounted for in the LDA pseudopotential. It would indeed be suspicious on the part of the WDA if the gaps for Ge and GaAs were improved with the use of LDA pseudopotentials.

5.4 Exchange-Correlation Holes in Si[110]

5.4.1 Comparison with the VMC Method

Analysing XC holes on a local scale provides a more stringent test of a density functional compared with examining global quantities such the total energy $E_{\rm XC}$, where errors may be averaged out by fortuitous cancellation. This is a valuable tool that is absent from most other functional types such the GGA and MGGA as explained in Chapter 2. For such functionals only spherically and system averaged holes can be determined [157, 158], however this procedure effectively smooths out the non-localities contained in the local hole.

In this section, XC holes are calculated with the various WDA pair-correlation functions and are compared with variational Monte Carlo data that has recently been made available. The VMC simulations, performed by Hood *et al.* [145], determined XC holes at specific positions in the [110]

plane of the valence density of Si (the core electrons are excluded due to pseudisation) that constitute different density environments. Attempts are made to re-create as best as possible the same conditions as in the VMC work, consequently the WDA holes were obtained using a density of Si determined from a self-consistent LDA calculation at the experimental lattice constant, as was the case with the VMC work of Hood *et al.*

5.4.2 Generation Procedure

5.4.2.1 The VMC Method

The (coupling-constant averaged) pair-correlation function $g_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')$ is the main quantity to determine in any method since the definition of the (coupling-constant averaged) hole (2.24) follows naturally from $g_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')$. In the VMC method, pair-correlation functions are determined through relations (2.25) and (2.27) once the many-body wavefunctions Ψ_λ at various values of the coupling constant λ have been obtained. In order to be able to plot $n_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')$ at any point within the Si crystal, Hood *et al.* devised an efficient way of calculating and storing the required six-dimensional $(\mathbf{r}\times\mathbf{r}')$ information by expanding $g^{\lambda}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')$ as a product of single-particle symmetrised plane waves $\phi_i(\mathbf{r})$,

$$g^{\lambda}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = \sum_{n,m} g^{\lambda}_{n,m}\phi_n(\mathbf{r})\phi^*_m(\mathbf{r}')\,, \tag{5.5}$$

with a converged plane-wave cutoff. The full space-group symmetry of the cell was taken advantage of so that the number of coefficients $g^{\lambda}_{n,m}$ was substantially reduced. The coupling-constant integration was performed numerically using five values of λ: $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ and 1. A more in-depth discussion of the method is presented in Ref. [159]. The VMC holes presented in this investigation were generated using this method.

5.4.2.2 The WDA Method

Exchange-correlation holes are generated very easily with the WDA, and proceed as follows: the density upon which the holes are to be calculated (in this case an LDA density of Si) is passed through a single WDA cycle with a given choice of function $G^{\mathrm{WDA}}[\mathbf{r}, \mathbf{r}'; \tilde{n}]$, this yields $\tilde{n}(\mathbf{r})$ and therefore $\alpha(\mathbf{r})$ and $\beta(\mathbf{r})$ at each point in the cell. For a specified location in the unit cell density, an electron is held fixed at $\mathbf{r}$ whilst the position of the reference electron $\mathbf{r}'$ is moved to all points within a pre-defined plane, and the corresponding hole surrounding the electron at $\mathbf{r}$ is constructed from $\alpha(\mathbf{r})$, $\beta(\mathbf{r})$ and $n(\mathbf{r}')$ using the relation,

$$n_{\mathrm{XC}}^{\mathrm{WDA}}(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}')\alpha(\mathbf{r})f(u)\,. \tag{5.6}$$

5.4.3 Results

Three positions of interest in the [110] plane of Si are analysed, which are shown in Fig. 5.5. The first point, a bond centre, is where the valence density of Si is at its greatest and also has moderate variations. The second point is at a pseudoatom centre where the density has strong variations because of interactions with the pseudised core-electrons. Finally, the region of lowest density, namely an interstitial site is examined.

5.4.3.1 Bond Centre

Due to the covalent nature of Si, the (valence) density is greatest at bond centres than anywhere else in the crystal, therefore an accurate description of the hole in this region is vital for the description of physical properties since this is where most of the XC energy is contained. The hole obtained from the VMC method at this point is shown in Fig. 5.6(a). The WDA exhibits a wide range of holes depending on the choice of the pair-correlation function. However, of the twelve functions studied, $f_1(u)$ is in closest agreement with the VMC data in terms of the depth and general shape of the hole, as observed

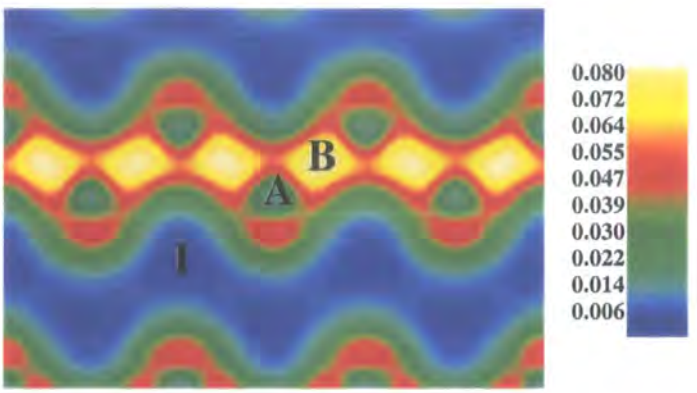

Figure 5.5: Valence density of Si in the [110] plane showing the three locations where exchange-correlation holes $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ are examined - a bond centre (labelled B), pseudoatom centre (labelled A) and at an interstitial site (labelled I).

from Fig. 5.6(b). Within each of the four groups, the depth, and therefore spatial extent of the hole, gradually decreases for each function. This is demonstrated for the functions of group 1 shown in Figs. 5.6(b) to (d). The function $f_6(u)$ yields the hole with the smallest on-top value (i.e. when $\mathbf{r} = \mathbf{r}'$) as shown in Fig. 5.6(e), differing from the VMC result by $\sim 50\%$. Fig. 5.6(f) shows the hole obtained from $f_7(u)$ which gives the deepest hole of all the functions - 30% greater than the VMC result.

With the exception of $f_{12}(u)$ which will be discussed later, the results obtained for the hole are directly linked to the accuracy of the lattice constants given in Table 5.2. For example, $f_6(u)$ and $f_7(u)$ give rise to the shortest and longest lattice constants respectively, and $f_1(u)$ yields the most accurate lattice constant in Table 5.2, apart from the function $f_2(u)$ which has the same mean absolute error.

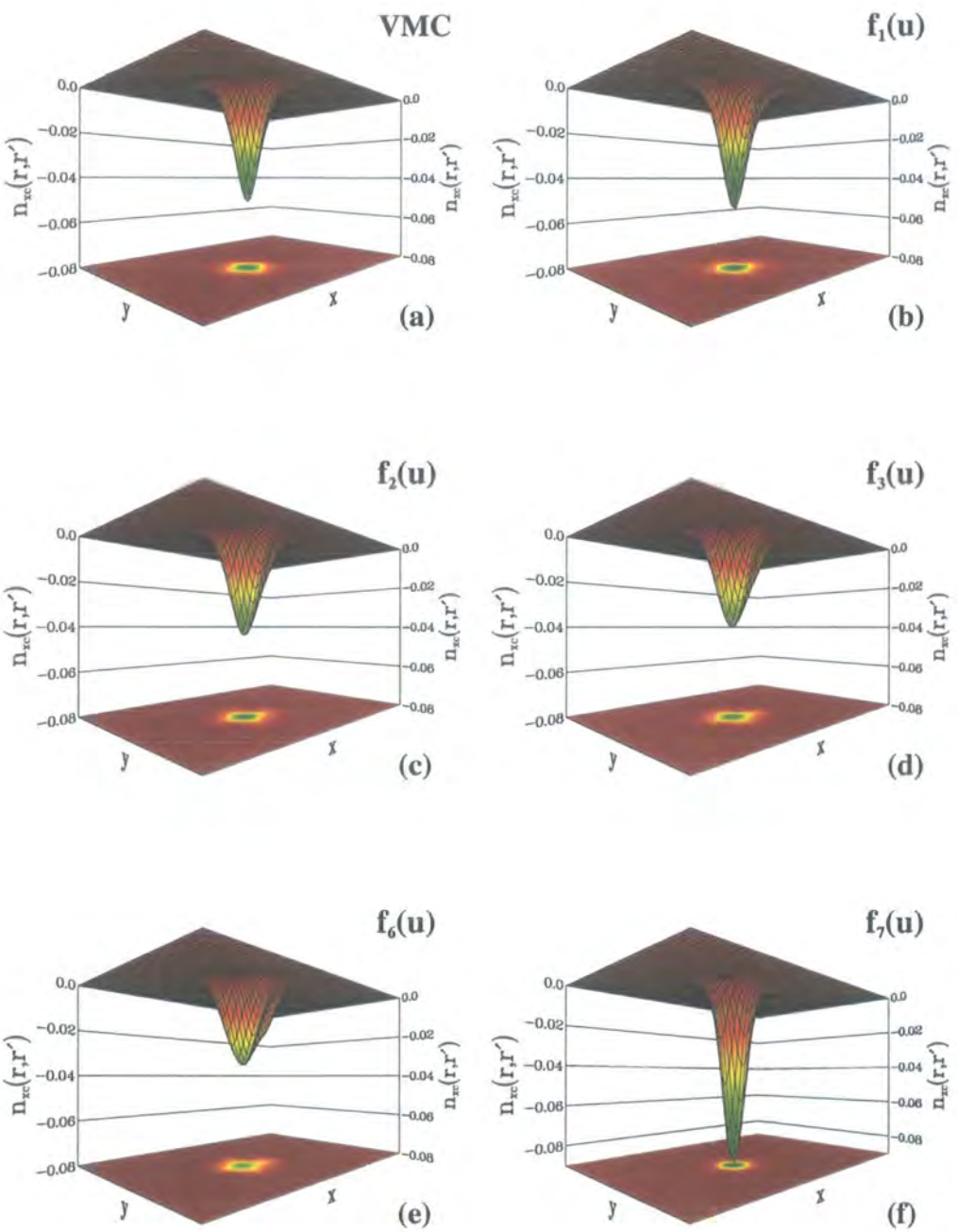

Figure 5.6: The XC hole $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ plotted in the [110] plane of Si for an electron located at a bond centre (position B in Fig. 5.5) using (a) VMC, and the WDA with model pair-correlation functions (b) $f_1(u)$, (c) $f_2(u)$, (d) $f_3(u)$, (e) $f_6(u)$ and (f) $f_7(u)$. The reference electron in situated at the minimum of the hole in all cases. Note the slight change in the scale in (f).

5.4.3.2 Pseudoatom Centre

The trend between the quality of the hole description and the structural data is also observed when the electron is situated at the pseudoatom centre (position A in Fig. 5.5). Fig. 5.7 shows the hole associated with an electron at this point given by (a) the VMC method and (b) $f_1(u)$, (c) $f_6(u)$, and (d) $f_7(u)$. The density is varying most rapidly at this position, consequently the holes display several minima that arise from the nearest bonds - the position of the reference electron being at the centre of these minima. All of the WDA functions produce deeper minima than the VMC method, but again, $f_1(u)$ is in closest agreement, yielding the shallowest of the WDA holes, $\sim$ 27% deeper than VMC. Even $f_6(u)$ which gave the shallowest hole at the bond centre yields a deeper hole than $f_1(u)$. The function $f_7(u)$ gives the largest discrepancy, being $\sim$ 75% deeper than the VMC result. Of course some of the discrepancy between the WDA and VMC data must be attributed to the differences between the densities used. The use of different pseudopotentials with different core radii will result in slight changes, especially for holes calculated at the pseudoatom site.

5.4.3.3 Interstitial Site

The hole at the interstitial site (position I in Fig. 5.5) will have less of an effect on structural properties because of the low density in this region, however it may be more important for conduction band properties because of the influence on the XC potential in this region. Nevertheless, the structure of the hole becomes much more interesting here since significant anisotropic effects emerge. The VMC hole, shown in Fig. 5.8(a), exhibits three non-local minima that surround the reference electron that is situated at the centre of the minima. This strong non-locality occurs because of the diffuse nature of the pair-correlation function which encompasses part of the rapidly increasing density from the three nearest atoms (see Fig. 5.5). This demonstrates

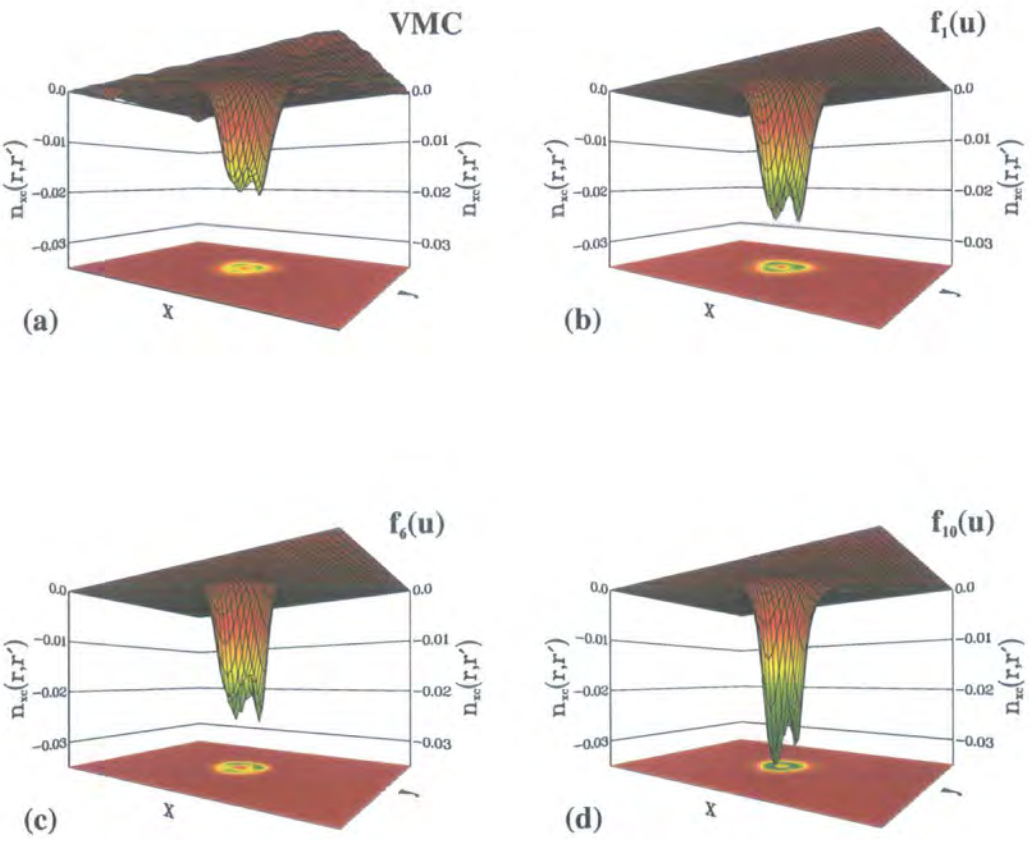

Figure 5.7: The XC hole $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ associated with an electron at the pseudoatom centre (labelled A in Fig. 5.5) for (a) VMC, and (b) the WDA function f_1. The xy plane is rotated by -135 degrees relative to Fig. 5.5.

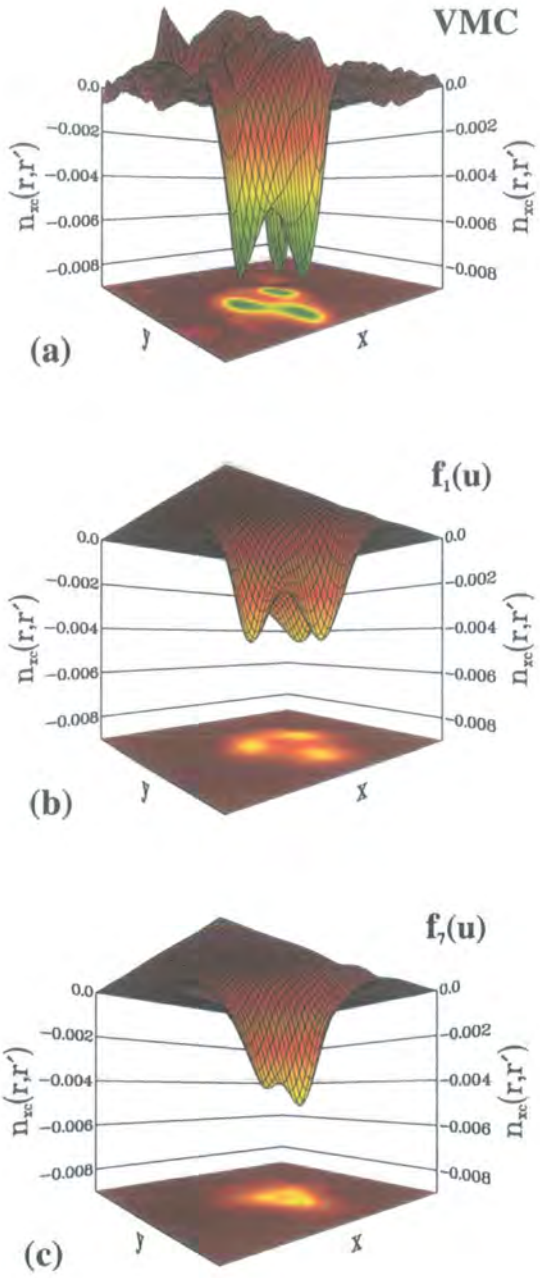

Figure 5.8: The exchange-correlation hole associated with an electron located at the interstitial site (position I in Fig. 5.5) obtained from (a) the VMC method, and the WDA with functions (b) $f_1(u)$, and (c) $f_7(u)$. The electron is located in the centre of the hole minima. The fluctuations in the VMC hole are statistical errors and are not part of the actual hole.

the significance of the surrounding density in the description of the hole. Representative examples of WDA holes are shown in Figs. 5.8(b) and 5.8(c) for the functions $f_1(u)$ and $f_7(u)$ respectively. With the exception of $f_7(u)$ and $f_{10}(u)$, the WDA functions are in qualitative agreement with the VMC data in that they reproduce the non-local minima, $f_1(u)$ being a typical example. However they each yield different values for the depth of the minima. The shallowest is given by $f_6(u)$ which is $\sim 33\%$ less than the VMC result. The XC holes given by $f_7(u)$ and $f_{10}(u)$ are fundamentally different from the rest of the WDA models, the compactness of these functions leads almost to the disappearance of the minima.

More understanding of these results can be gained from the pair-correlation function, $g_{\rm XC}(\mathbf{r},\mathbf{r}')$, that gives rise to the XC hole. Fig. 5.9 shows $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ calculated at the interstitial point for (a) the VMC method, (b) $f_1(u)$ and (c) $f_7(u)$. It is clear that $f_1(u)$ possesses a similar qualitative shape as the VMC result, although the on-top value $g_{\rm XC}(\mathbf{r},\mathbf{r})$, is slighter smaller. In contrast, $f_7(u)$ yields a very different result in comparison to $f_1(u)$ and the VMC method. Indeed, a serious deficiency occurs with $f_7(u)$, and similarly for $f_{10}(u)$ although to a lesser extent, in that the pair-correlation function becomes negative in the vicinity of the electron, $\mathbf{r}' \rightarrow \mathbf{r}$, which is completely unphysical. The tendency for $f_7(u)$ and $f_{10}(u)$ pair-correlation functions to become negative at short range distances was hinted at earlier for the case of the homogeneous electron gas discussed in Sec. 5.2. This problem has also been encountered in other studies of the pair-correlation function [160]. The unusually large band gaps obtained with $f_7(u)$ and $f_{10}(u)$ may be related to the unphysical behaviour displayed by these long-ranged pair-correlation functions.

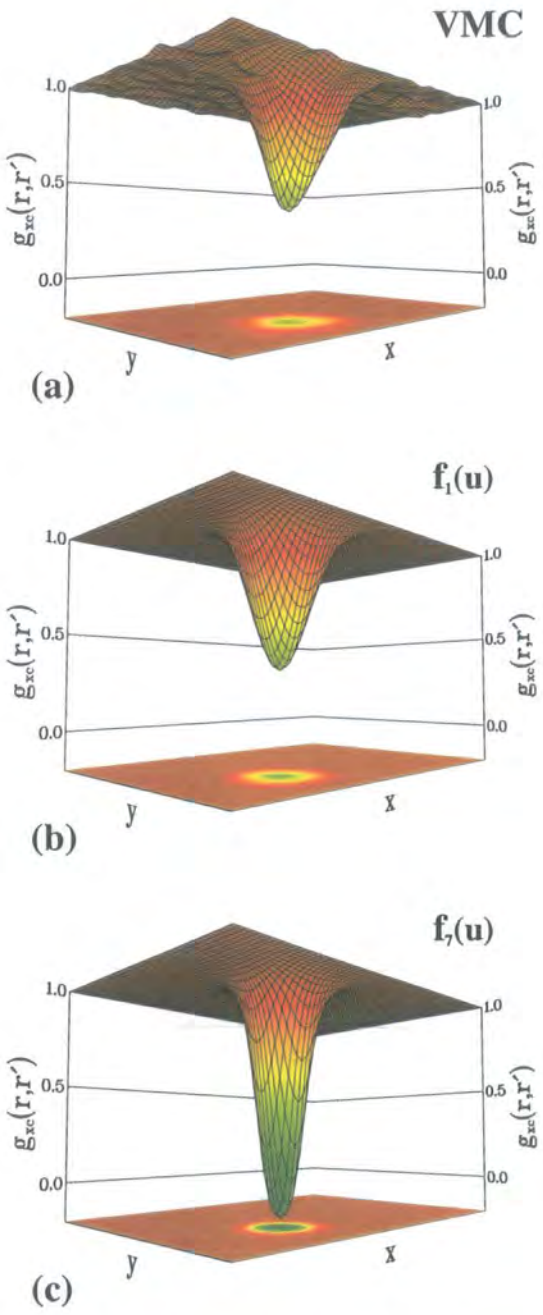

Figure 5.9: The pair-correlation function $g_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ associated with an electron located at the interstitial site (labelled I in Fig. 5.5) obtained from (a) the VMC method, and the WDA with functions (b) f_1, and (c) f_7.

5.4.4 Discussion

An interesting result comes from the WDA function $f_{12}(u)$ which from the structural data given in Table 5.2, appears to be the best model pair-correlation function of the ones tested. However, this function provides a worse description of the XC hole at the bond and pseudoatom centres in comparison with $f_1(u)$. The corresponding holes for $f_{12}(u)$ are shown in Fig. 5.10 for an electron located at the bond and pseudoatom centres. The on-top value in the

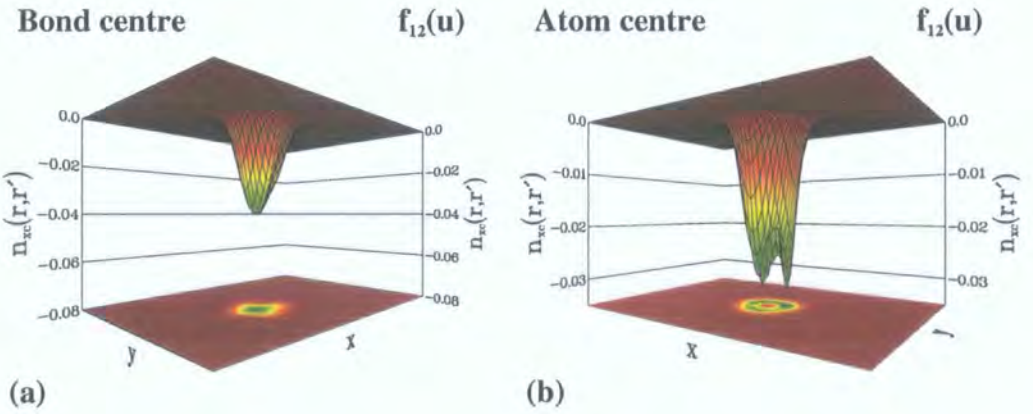

Figure 5.10: The exchange-correlation hole $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ calculated with the WDA using function $f_{12}(u)$, for an electron located at (a) the bond and (b) pseudoatom centres, in Si[110].

bond centre is $\sim 20\%$ less than the VMC method, whilst at the pseudoatom centre the hole is $\sim 55\%$ deeper than the VMC data. Therefore, the analysis of XC holes reveals that this particular function relies on a fortuitous cancellation of errors in order to produce the excellent structural properties. This highlights the caution that should be undertaken when assessing functionals based solely on global quantities such as structural and energetic data, since opposing errors can easily cancel, which will result in wrong deductions. Analysing XC holes in this way provides very useful insight into a functional.

5.5 Development

5.5.1 Existing Problems

The twelve functions considered previously contain certain deficiencies that potentially limit their performance. Firstly, all of the functions become negative in the short-range limit ($r \rightarrow 0$) for sufficiently low densities, which violates the positive definite character of the exact pair-correlation function. The degree to which this failure manifests depends on the particular function and in most cases was not important for the range of densities encountered in bulk Si, the functions $f_7(u)$ and $f_{10}(u)$ being the obvious exception. Also, in the high density limit ($r_s \rightarrow 0$), all of the functions give rise to $g_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}') > 0.5$ which is again a violation of an exact condition. Another limitation of the functions is that the Kimball cusp condition (5.4) is not satisfied. It is therefore important to address these problems if greater accuracy and universality is to be achieved with the WDA.

This section presents a more flexible model pair-correlation function model that obeys the Kimball cusp condition, denoted $G^{\mathrm{Cusp}}[\mathbf{r}, \mathbf{r}'; \tilde{n}(\mathbf{r})]$, and also attempts to resolve the issue of constraining $g_{\mathrm{xc}}(\mathbf{r}, \mathbf{r})$ to the specified range $[0, 0.5]$. The object of this exercise is mainly just to highlight some developmental points.

5.5.2 A New Model Function

The philosophy behind the development undertaken here was to base the new model on the simple Gaussian function $f_1(u)$, as this was the most promising of the twelve functions studied previously, and then incorporate the Kimball cusp condition by altering only the short-range character of this function. The resulting model, called $G^{\mathrm{Cusp}}[\mathbf{r}, \mathbf{r}'; \tilde{n}]$, consists of two terms,

$$G^{\mathrm{Cusp}}[\mathbf{r}, \mathbf{r}'; \tilde{n}] = G^{a}[\mathbf{r}, \mathbf{r}'; \tilde{n}] + G^{b}_{\kappa,m}[\mathbf{r}, \mathbf{r}'; \tilde{n}]\,. \tag{5.7}$$

where the first term is the original Gaussian model,

$$G^a[\mathbf{r},\mathbf{r}';\tilde{n}] = \alpha(\tilde{n})\,\mathrm{e}^{-\left(\frac{|\mathbf{r}-\mathbf{r}'|}{\beta(\tilde{n})}\right)^2}, \tag{5.8}$$

and the second incorporates the cusp condition,

$$G^b_{\kappa,m}[\mathbf{r},\mathbf{r}';\tilde{n}] = (\alpha(\tilde{n})+1)|\,\mathbf{r}-\mathbf{r}'|\,\mathrm{e}^{-\left(\frac{|\mathbf{r}-\mathbf{r}'|}{\kappa\beta(\tilde{n})}\right)^m}. \tag{5.9}$$

Again, shorthand notation is used such that $\tilde{n} = \tilde{n}(\mathbf{r})$. The role of the parameters m and κ is to constrain the influence of $G^b_{\kappa,m}[\mathbf{r},\mathbf{r}';\tilde{n}]$ to short-ranged interactions only - m adjusts the general shape of the modification, while κ directly alters its range, so that the behaviour of $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$ is left unchanged for large inter-electron separations. The range of influence of $G^b_{\kappa,m}[\mathbf{r},\mathbf{r}';\tilde{n}]$ is directly proportional to κ, so when $\kappa = 0$, the range of the function is also zero, and the model reverts back to $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$. Different values for m and κ will be investigated in Sec. 5.5.4 - the general strategy for determining their values is to vary them in such a way as to give on-top pair-correlation values, $g_{\mathrm{XC}}(\mathbf{r},\mathbf{r})$, that stay within the range $[0, 0.5]$.

Except in the obvious case where $\kappa = 0$, the new model satisfies the Kimball cusp condition for all choices of m and κ, i.e.

$$\left.\frac{\partial g^{\mathrm{Cusp}}_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')}{\partial r}\right|_{\mathbf{r}=\mathbf{r}'} = \left. g^{\mathrm{Cusp}}_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')\right|_{\mathbf{r}=\mathbf{r}'} = \alpha(\tilde{n})+1\,, \tag{5.10}$$

where $g^{\mathrm{Cusp}}_{\mathrm{XC}}(\mathbf{r},\mathbf{r}') = G^{\mathrm{Cusp}}[\mathbf{r},\mathbf{r}';\tilde{n}]+1$. The newly proposed model is probably the simplest way to incorporate the Kimball cusp condition within the existing WDA framework, and can be implemented within the original computer code with only a few minor adjustments. These points are discussed next.

5.5.3 Implementation Details

Re-writing (5.7) in the usual fashion by substituting $u = |\,\mathbf{r}-\mathbf{r}'\,|/\beta(\tilde{n})$, yields,

$$G^{\mathrm{Cusp}}[u,\tilde{n}] \;=\; \alpha(\tilde{n})f^a(u) \;+\; (\alpha(\tilde{n})+1)\beta(\tilde{n})f^b(u)\,, \tag{5.11}$$

where $f^a(u) = \mathrm{e}^{-u^2}$ and $f^b(u) = u\mathrm{e}^{-(u/\kappa)^m}$. As a result, the reciprocal-space forms of the XC energy density and sum rule are written in a similar fashion as before,

$$\varepsilon_{\mathrm{XC}}(\mathbf{r}) = \frac{1}{2}\,\alpha(\tilde{n})\,\beta^2(\tilde{n}) \sum_{\mathbf{G}} n(\mathbf{G})\,\mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} H_1\,, \tag{5.12}$$

$$-1 = \alpha(\tilde{n})\,\beta^3(\tilde{n}) \sum_{\mathbf{G}} n(\mathbf{G})\,\mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} H_2\,, \tag{5.13}$$

except the functions H_1 and H_2, are given by,

$$H_p = F_p^a(q) + \left(1 + \frac{1}{\alpha(\tilde{n})}\right) \beta(\tilde{n})\, F_p^b(q)\,, \tag{5.14}$$

where $p = 1, 2$, and $q = \beta(\tilde{n})\,|\mathbf{G}|$ as usual, and the functions $F_p^a(q)$ and $F_p^b(q)$ are defined as,

$$F_p^\sigma(q) = \frac{4\pi}{q} \int_0^\infty u^{p-1}\, f^\sigma(u) \sin(qu)\, du\,, \tag{5.15}$$

where $\sigma = a, b$. For a homogeneous electron gas, relations (5.12) and (5.13) can be rearranged to give the following definition of $\alpha(\tilde{n})$ as,

$$\alpha(\tilde{n}) = -\frac{1 + \tilde{n}\beta^4(\tilde{n})\, I_2^b}{\tilde{n}\,\beta^3(\tilde{n})\,(I_2^a + \beta(\tilde{n})\, I_2^b)}\,. \tag{5.16}$$

The value of $\beta(\tilde{n})$ is obtained by substituting (5.16) into (5.12) and rearranging. The result is a fourth order polynomial with the form,

$$a_4\beta^4(\tilde{n}) + a_2\beta^2(\tilde{n}) + a_1\beta(\tilde{n}) + a_0 = 0\,, \tag{5.17}$$

where the coefficients are,

$$a_4 = \tilde{n}\,(I_1^a I_2^b - I_2^a I_1^b)\,, \tag{5.18}$$

$$a_2 = 2 I_2^b \varepsilon_{\mathrm{XC}}^{\mathrm{LDA}}(\tilde{n})\,, \tag{5.19}$$

$$a_1 = I_1^b + 2\, I_2^a \varepsilon_{\mathrm{XC}}^{\mathrm{LDA}}(\tilde{n})\,, \tag{5.20}$$

$$a_0 = I_1^a\,. \tag{5.21}$$

The constants I_p^σ are defined as,

$$I_p^\sigma = 4\pi \int_0^\infty u^p f^\sigma(u)\, du\,. \tag{5.22}$$

Again, the potential contains three terms as in the original case,

$$v_{\rm XC}(\mathbf{G}) = v_1(\mathbf{G}) + v_2(\mathbf{G}) + v_3(\mathbf{G}) \tag{5.23}$$

with $v_1(\mathbf{G}) = \varepsilon_{\rm XC}(\mathbf{G})$ and,

$$\begin{aligned} v_2(\mathbf{G}) &= \frac{2\pi}{\Omega}\int n(\mathbf{r}')\,\alpha(\mathbf{r}')\,\beta^2(\mathbf{r}')\,\mathrm{e}^{-i\mathbf{G}\cdot\mathbf{r}'}\,H_1\,d\mathbf{r}'\,, \\ v_3(\mathbf{G}) &= -\frac{\pi}{\Omega}\int n(\mathbf{r}')\,\alpha(\mathbf{r}')\,\beta^3(\mathbf{r}')\,\frac{h_1(\mathbf{r}')}{h_2(\mathbf{r}')}\,\mathrm{e}^{-i\mathbf{G}\cdot\mathbf{r}'}\,H_2\,d\mathbf{r}'\,. \end{aligned} \tag{5.24}$$

The functions $H_1(\mathbf{r})$ and $H_2(\mathbf{r})$ are given by:

$$\begin{aligned} h_p(\mathbf{r}) &= \beta(\tilde{n})\sum_{\mathbf{G}} n(\mathbf{G})\,\mathrm{e}^{i\mathbf{G}\cdot\mathbf{r}} \\ &\quad\times\left[\beta(\tilde{n})\frac{\partial\alpha(\tilde{n})}{\partial\tilde{n}}\left\{F_p^a(q) + \beta(\tilde{n})\,F_p^b(q)\right\} - \alpha(\tilde{n})\,\frac{\partial\beta(\tilde{n})}{\partial\tilde{n}}\,H_p'\right], \end{aligned} \tag{5.25}$$

where the quantities H_p' are given by,

$$H_p' = F_p'^a(q) + \left(1 + \frac{1}{\alpha(\tilde{n})}\right)\beta(\tilde{n})\,F_p'^b(q)\,, \tag{5.26}$$

and

$$F_p'^a(q) = \frac{8\pi}{q}\int_0^\infty u^{p+1}\,f^a(u)\sin(qu)\,du\,, \tag{5.27}$$

$$F_p'^b(q) = \frac{4\pi}{q}\frac{m}{\kappa^m}\int_0^\infty u^{m+p-1}\,f^b(u)\sin(qu)\,du\,. \tag{5.28}$$

The derivative $\partial\beta(\tilde{n})/\partial\tilde{n}$ can be obtained by implicit differentiation of relation (5.17), yielding,

$$\frac{\partial\beta(\tilde{n})}{\partial\tilde{n}} = \frac{\beta^4(\tilde{n})(a_4/\tilde{n}) - 2\beta(\tilde{n})(I_2^a + \beta(\tilde{n})I_2^b)(d\varepsilon_{\rm XC}^{\rm LDA}(\tilde{n})/d\tilde{n})}{4\beta(\tilde{n})(a_4\beta^2(\tilde{n}) + (a_2/2)) + a_1} \tag{5.29}$$

where $d\varepsilon_{\rm XC}^{\rm LDA}(\tilde{n})/d\tilde{n} = (v_{\rm XC}^{\rm LDA}(\tilde{n}) - \varepsilon_{\rm XC}^{\rm LDA}(\tilde{n}))/\tilde{n}$. $\partial\alpha(\tilde{n})/\partial\tilde{n}$ can be obtained from relation (5.16),

$$\frac{\partial\alpha(\tilde{n})}{\partial\tilde{n}} = \frac{1}{\tilde{n}^2\beta^3(\tilde{n})(I_2^a + I_2^b\beta(\tilde{n}))} + \frac{\partial\alpha(\tilde{n})}{\partial\beta(\tilde{n})}\frac{\partial\beta(\tilde{n})}{\partial\tilde{n}}\,, \tag{5.30}$$

and,

$$\frac{\partial\alpha(\tilde{n})}{\partial\beta(\tilde{n})} = \frac{3I_1^b + 4I_2^b\beta(\tilde{n}) - \tilde{n}I_2^aI_2^b\beta(\tilde{n})}{\tilde{n}\beta^4(\tilde{n})(I_2^a + I_2^b\beta(\tilde{n}))^2}\,. \tag{5.31}$$

The computational expense of the new implementation is only slightly greater than the original scheme. The extra time is mainly devoted to the solution of the fourth order polynomial (5.17) which must be solved each time the value of $\tilde{n}$ is updated. The new implementation is however more general as it reverts back to the original scheme when the value of κ is set to zero.

5.5.4 Results for the Homogeneous Electron Gas

As a first brief test, the pair-correlation function $g_{\rm XC}(\mathbf{r},\mathbf{r}')$ obtained from the new model is examined in the case of the homogeneous electron gas, for densities ranging from high ($r_s = 0.1$), to low ($r_s = 10.0$) values.

Two versions of the new model are considered which employ the values $m = 2$, and $m = 4$, and are labelled $G_1^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ and $G_2^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ respectively. As mentioned in Sec. 5.5.1, another objective of the new model, apart from incorporating the cusp condition, is to tackle the problem of the on-top pair-correlation function violating the exact range of values $[0, 0.5]$. Therefore in each of the two cases considered, the value of κ was set by attempting to keep the on-top values within this specified range. The optimal values, $\kappa = 0.75$ and $\kappa = 1.08$, were obtained for $G_1^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ and $G_2^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ respectively. Table 5.6 compares the on-top value of the homogeneous electron gas pair-correlation function $g_{\rm XC}(r = 0)$, calculated for the range of densities $0.1 \leq r_s \leq 10.0$, with the original Gaussian model $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$, and the two new versions. It is observed from these results that although the new models still yield $g_{\rm XC}(r = 0)$ greater than 0.5 in the high density case, and lower then 0 in the low density regime, they are an improvement over $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$. For $r_s = 0.1$, the on-top values are 0.564, 0.538 and 0.516 for $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$, $G_1^{\rm cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$, and $G_2^{\rm cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ respectively, whereas for $r_s = 10.0$, the values are -0.118, -0.026 and -0.015 respectively. The function $G_2^{\rm cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ is therefore the most successful in this respect.

Table 5.6: The on-top value of the pair-correlation function, $g_{\rm XC}(\mathbf{r},\mathbf{r})$, for the homogeneous electron gas at various r_s values. Results are presented for the original Gaussian pair-correlation function that violates the cusp condition $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$, and two versions of the new model that satisfy the cusp condition, $G_1^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ with $m=2,\kappa=0.75$, and $G_2^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$ with $m=4,\kappa=1.08$.

r_s	G^a	$G_1^{\rm Cusp}$	$G_2^{\rm Cusp}$
0.1	0.564	0.538	0.516
0.2	0.541	0.494	0.457
0.3	0.522	0.457	0.410
0.4	0.504	0.425	0.371
0.5	0.488	0.397	0.338
0.6	0.473	0.372	0.310
0.8	0.444	0.328	0.263
1.0	0.419	0.291	0.226
2.0	0.309	0.169	0.119
3.0	0.225	0.102	0.068
4.0	0.155	0.060	0.039
5.0	0.096	0.033	0.021
6.0	0.044	0.013	0.008
8.0	−0.045	−0.011	−0.007
10.0	−0.118	−0.026	−0.015

The effect on the overall pair-correlation function is observed in Fig. 5.11 which shows $g_{\rm XC}(r)$ calculated using $G^a[\mathbf{r},\mathbf{r}';\tilde{n}]$ and $G_1^{\rm Cusp}[\mathbf{r},\mathbf{r}';\tilde{n}]$, for several r_s values. It is clear that while both models demonstrate different behaviour near the cusp, as expected, they are very similar as r increases, as intended. However, an unfortunate feature of the new models is that in the low density regime, they yield a negative gradient at $r=0$, which is a consequence of satisfying the cusp condition (5.10) when $g_{\rm XC}(r)$ goes negative. An example of this is shown for the $r_s=10.0$ case. This unphysical characteristic can only be eliminated by satisfying the non-negativity constraint for all r_s.

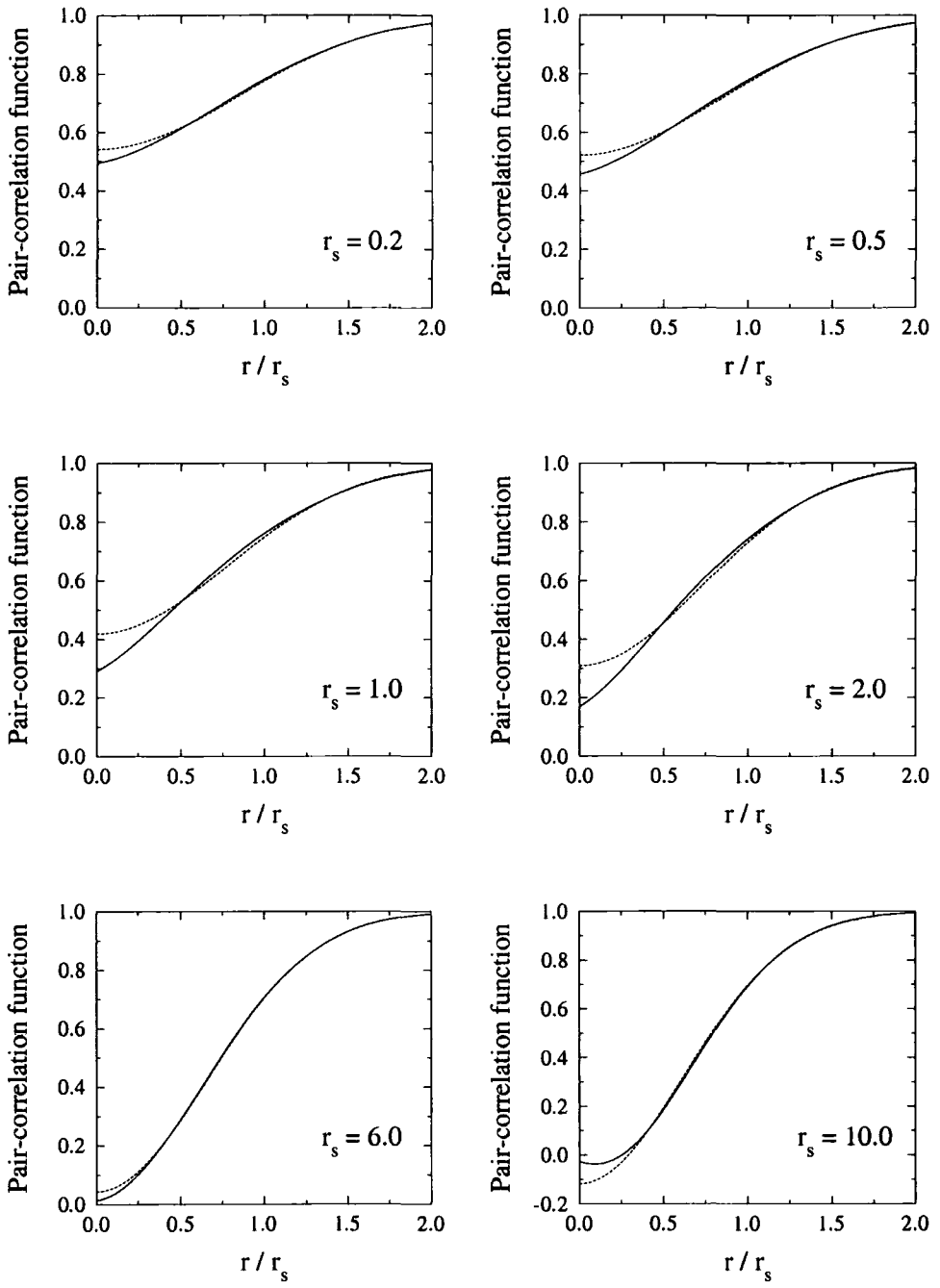

Figure 5.11: The pair-correlation function $g_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$ calculated for the homogeneous electron gas at various r_s values, determined using the new model $G_1^{\mathrm{Cusp}}[\mathbf{r}, \mathbf{r}'; \tilde{n}]$, that incorporates the Kimball cusp condition (solid lines), and the original Gaussian model $G^a[\mathbf{r}, \mathbf{r}'; \tilde{n}]$ (dotted lines).

5.6 Summary

This chapter demonstrated that the form of the analytic function used to model the coupling-constant averaged pair-correlation function in the WDA is crucially important. WDA properties are indeed sensitive to the type of model employed. This was observed by the wide range of lattice constants and bulk moduli calculated for Si using twelve example models. The extremely weak nature of the bonding reported by Charlesworth in Ref. [150] for the WDA using the same model functions, was not observed in this study. In great contrast to Charlesworth, some of the models yielded lattice constants shorter than experiment and most were closer to experiment than the LDA. Hopefully the new results presented here, which are substantiated by the comparison of XC holes with the variational Monte Carlo method, will stimulate further research into model pair-correlation functions for use in the WDA.

Although the specific choice of model functions provided informative insights into general trends in the structural properties and demonstrated close links with the corresponding XC holes, they all possess certain deficiencies. One particular fault is the violation of the non-negativity constraint. The models $f_7(u)$ and $f_{10}(u)$ violated this condition most severely of the twelve models considered, and also produced the greatest errors in the structural properties and holes.

In order to tackle such problems a new prescription was proposed that satisfies an extra exact constraint known as the Kimball cusp condition. In addition, the new prescription attempted to resolve the fact that the on-top pair-correlation function should be in the range $[0, 0.5]$, by including two extra parameters, namely m and κ. Although this objective was not achieved precisely, a definite improvement over the original formulation was attained when examining the homogeneous electron gas. The inclusion of the additional parameters does not invalidate the non-empirical nature of

the functional, since the parameters were invoked in order to satisfy exact constraints. A model function that satisfies the on-top condition for the homogeneous electron gas, for a broad range of densities, must first be devised before applied to real inhomogeneous systems. Hopefully this approach to testing the models will prove universally applicable, i.e. for the majority of real physical systems.

Another direction that could be pursued in the development of model pair-correlation functions is to devise models for exchange and correlation separately. Incorporating coupling-constant (λ) dependence may also be useful due to the greater amount of near-exact data that could be utilised - QMC simulations are again a good source of this type of data since they must calculate the pair-correlation function at several values of λ in order to determine (λ-averaged) density functional exchange-correlation quantities.

Chapter 6

Model Electron Gas Systems

6.1 Introduction

The reason for developing functionals is so that the properties of matter can be predicted more accurately, hence the determination of real atomic, molecular and solid state quantities represents an important and necessary test of any new functional. However, to obtain a true assessment of an exchange-correlation functional that is implemented within a plane-wave formalism, it is essential to use consistent pseudopotentials in the calculations - as demonstrated in Sec. 3.3.6. The generation of WDA pseudopotentials may be complicated by the need to include shell-partitioning techniques [85, 147], whereby the core intra-shell exchange-correlation interactions are accounted for by the WDA, while the inter-shell regions are described with the LDA. However as yet, no attempt has been made to do this. While the self-consistent WDA calculations performed in the previous chapter demonstrated general trends, they are by no means conclusive results. So an extensive study of solid state systems will provide only limited insight into the WDA whilst inconsistent pseudopotentials are used. Although the development of WDA pseudopotentials is of utmost importance to the WDA method as a whole, they are not the focus of this thesis and so will not be

discussed further.

This chapter focuses on analysing the WDA in various inhomogeneous electron gas systems. The reason for choosing such model systems is because the electron gas does not contain ions, so the pseudopotential issue is automatically resolved. Also, the density can be completely controlled, so exchange-correlation interactions can be explored in a range of pre-determined density regimes. There exist a plethora of electron gas systems that can be examined quite easily, however the particular examples chosen here were stimulated by a recent quantum Monte Carlo study performed by Nekovee *et al.* [161], which applied a strong one-dimensional cosinusoidal potential to a three-dimensional electron gas. These systems will be re-examined using the WDA, along with other extensions including the quasi-2D limit of the electron gas, and strong isotropic confinement in three dimensions.

6.2 The Cosine-Wave Electron Gas

The homogeneous electron gas, also known as the jellium model, consists of N electrons enclosed in a box of volume, V, that is periodically repeated in space, and has a background of positive neutralising charge. In this section, three systems are examined in which a cosinusoidal perturbation, $v_{\text{ext}}(\mathbf{r})$, is applied along one direction of a three dimensional electron gas, with the form:

$$v_{\text{ext}}(\mathbf{r}) = v_{\text{q}} \cos(\mathbf{q} \cdot \mathbf{r}) \,, \tag{6.1}$$

where the amplitude v_q and wavevector $\mathbf{q}$ control the strength of the perturbation along a single dimension, leaving the system homogeneous in the other two dimensions. To be consistent with the VMC calculations in Ref. [161], the densities are generated from self-consistent DFT calculations using the

LDA, so the Kohn-Sham one-electron Hamiltonian $\hat{H}_{\rm KS}$ has the form,

$$\hat{H}_{\rm KS} = -\frac{1}{2}\nabla^2 + v_{\rm ext}(\mathbf{r}) + v_{\rm H}(\mathbf{r}) + v_{\rm XC}^{\rm LDA}(\mathbf{r})\,. \tag{6.2}$$

This procedure gives rise to density profiles that are approximately sinusoidal along the direction of inhomogeneity, which is labelled the y-direction. The systems are characterised by the value of $\mathbf{q}$ since $v_{\rm q}$ remains constant in all three cases. Each system has the same average density, with $r_s = 2.0a_0$, and the cells are designed such that they admit two, three and four periods of the cosine potential within the cell length. As a result, the density profiles exhibit the corresponding number of peaks along the direction of inhomogeneity. Along with the number of electrons, N, this defines the three values of the wavevector $\mathbf{q}$, which are given in terms of the Fermi energy $k_{\rm F}^0$. The particular values of $\mathbf{q}$ and $v_{\rm q}$ lead to strong variations in the density on the scale of the local Fermi wavelength, $\lambda_{\rm F}^0 = 2\pi/k_{\rm F}^0$.

6.2.1 Overview of a VMC Study

6.2.1.1 Details of the Calculations

Nekovee *et al.* used a method known as fixed-density variance minimisation to perform the VMC calculations, whereby the simulations are performed on a fixed density $n(\mathbf{r})$ obtained from an LDA calculation. Details of the procedure are given in Ref. [162]. Nekovee *et al.* defined the exchange-correlation energy density $e_{\rm XC}([n(\mathbf{r})],\mathbf{r})$ to be,

$$e_{\rm XC}([n(\mathbf{r})],\mathbf{r}) = \frac{n(\mathbf{r})}{2}\int \frac{n_{\rm XC}^{\rm VMC}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\,d\mathbf{r}'\,, \tag{6.3}$$

which differs from the usual definition (2.11) through the inclusion of $n(\mathbf{r})$. The total exchange-correlation energy iis therefore,

$$E_{\rm XC}[n(\mathbf{r})] = \int e_{\rm XC}([n(\mathbf{r})],\mathbf{r})\,d\mathbf{r}\,. \tag{6.4}$$

The VMC XC hole $n_{\rm XC}^{\rm VMC}(\mathbf{r},\mathbf{r}')$ used in (6.3), was determined from a numerical coupling-constant integration over six values of λ: 0, 0.2, 0.4, 0.6, 0.8, 1.0.

All of the VMC calculations were performed on a CRAY T3E supercomputer. The evaluation of each λ-dependent many-body wavefunction Ψ_λ took approximately 384 CPU hours, so that a total of 1920 CPU hours [163] were required to calculate the λ-averaged quantities for each of the three systems. (Note that only five simulations are actually performed since the $\lambda = 0$ wavefunction is just a slater determinant of single-particle orbitals which is already known from the input density $n(\mathbf{r})$). A further 350 CPU hours were required to collect statistics associated with each λ, which amounts to 2100 CPU hours for each system. So the total amount of time needed for the VMC calculations for each system was approximately 4020 CPU hours [163], using a CRAY T3E.

Nekovee *et al.* also performed VMC simulations on the homogeneous electron gas (HEG) at different densities in order to eliminate finite-size errors. As a result, the HEG coefficients contained in the correlation energy were re-parametrised, so the particular LDA and GGA forms used by Nekovee *et al.* are different from the conventional forms. Consequently, this will contribute to slight differences in the quoted LDA and GGA values.

6.2.1.2 Principal VMC Findings

Two particularly striking results emerged from the VMC simulations. The first concerns the similarity of the point-wise difference between the LDA and the VMC energy densities, $\Delta e_{\rm xc}(y)$, and the Laplacian of the density $\nabla^2 n(\mathbf{r})$, in all of the systems. Fig. 6.1, taken from Ref. [161], shows how close the relationship is between the two quantities, in terms of the shape, sign and magnitude. Since this effectively represents the non-locality missing from the LDA they proposed that $\nabla^2 n(\mathbf{r})$ should be included in semi-local density functionals.

The second finding regards the degree of non-locality exhibited by the exchange-correlation holes in these systems, particularly at the density minima. Fig. 6.2 shows the VMC and LDA holes near to a density maximum,

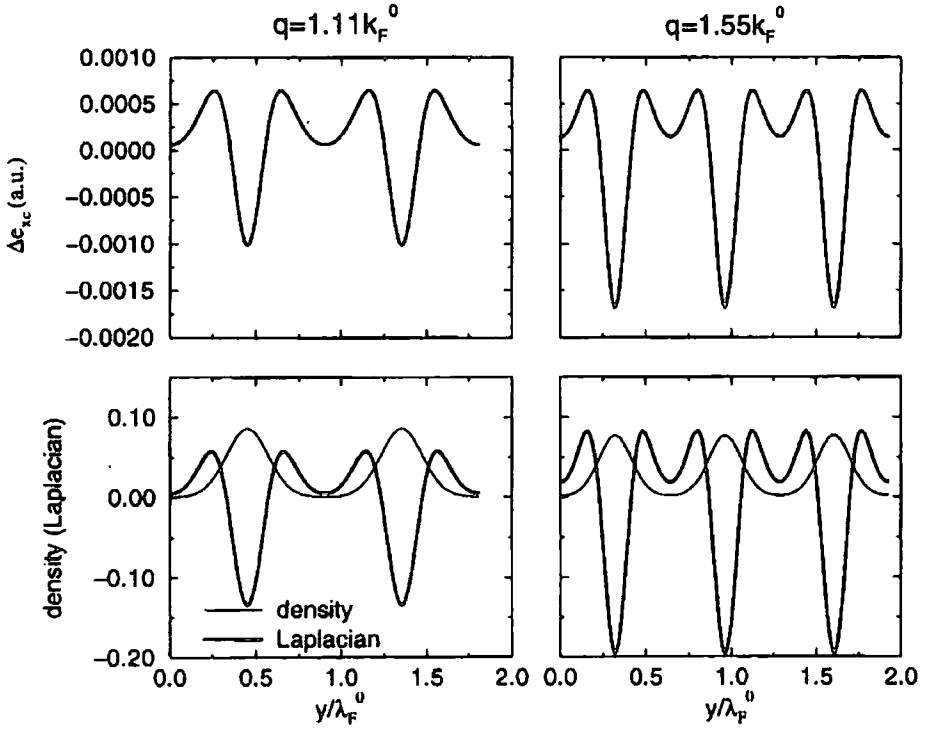

Figure 6.1: The difference between the LDA and VMC energy densities, $\Delta e_{\rm XC}(y) = e_{\rm XC}^{\rm LDA}(y) - e_{\rm XC}^{\rm VMC}(y)$, compared with $n(\mathbf{r})$ and $\nabla^2 n(\mathbf{r})$, for the cosine-wave electron gas systems with wavevector $q = 1.11\,k_{\rm F}^0$ and $q = 1.55\,k_{\rm F}^0$, calculated by Nekovee *et al.* in Ref. [161].

and at a minimum. In contrast to the LDA holes which are always spherical, the VMC holes are contracted in the direction of density inhomogeneity, and at the minimum the VMC hole displays two distinct minima that are widely separated. It is evident from this figure that the LDA is unable to demonstrate such non-local effects.

6.2.2 Details of the Calculations

The details of the systems studied here are designed to follow the work of Nekovee *et al.* as closely as possible, so that accurate comparisons can be

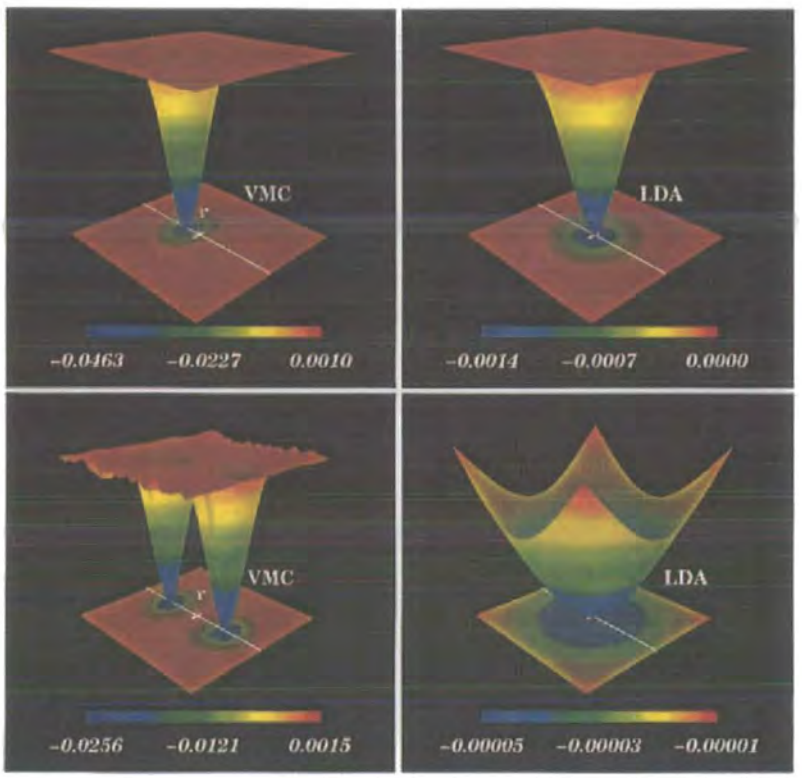

Figure 6.2: VMC and LDA exchange-correlation holes, $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$, generated by Nekovee *et al.* for the $q = 1.11\, k_{\mathrm{F}}^0$ cosine-wave system. The white line illustrates the direction of inhomogeneity. The top and bottom panels are for an electron situated near to a density maximum, and at a density minimum, respectively. Figures are taken from Ref. [161].

made between the WDA and VMC results. To achieve this, the densities were determined in the same fashion, i.e. through self-consistent DFT-LDA calculations. The same average density is used, i.e. $r_s = 2.0a_0$, which as a consequence defines the ratio N/V,

$$\frac{N}{V} = \frac{3}{4\pi r_s^3}, \tag{6.5}$$

and also the average Fermi wavevector k_{F}^0,

$$k_{\mathrm{F}}^0 = \left(\frac{9\pi}{4}\right)^{1/3} \frac{1}{r_s}. \tag{6.6}$$

A cubic cell is used for all three systems, therefore the wavevector, which has components $\mathbf{q} = (q, 0, 0)$, is related to V through the relation,

$$q = \frac{2\pi P}{V^{1/3}}, \tag{6.7}$$

where P is the number of periods of the cosine potential admitted across the length of the cell in each of the three cases. Using the above relations, to obtain wavevectors as close as possible as those employed by Nekovee *et al.*, each system must contain $N = 48$, 60 and 52 electrons. This leads to wavevectors of size $q = 1.12\,k_{\mathrm{F}}^0$, $1.56\,k_{\mathrm{F}}^0$ and $2.18\,k_{\mathrm{F}}^0$ respectively. The same amplitude is used in the calculations, corresponding to $v_{\mathrm{q}} = 2.08\,\varepsilon_{\mathrm{F}}^0$, where $\varepsilon_{\mathrm{F}}^0$ is the average Fermi energy.

The Gaussian pair-correlation $f_1(u)$ given in Table 5.1 is used in the WDA calculations unless otherwise stated. This model was chosen because of the promising results obtained in the study of Si in Chapter 5. The GGA calculations refer to the PBE [63] functional.

To give a perspective on the computational times invloved in relation to the VMC simulations, on a Compaq Alpha XP1000 667MHz computer, the WDA calculations require less than two hours for each of the three systems using a mesh of size $30\times2\times2$ for the WDA spatial interpolations, and 600 grid points for the weighted density interpolations. This should be compared with 4020 CPU hours on a CRAY T3E supercomputer for the VMC calculations, as mentioned in Sec. 6.2.1.1.

6.2.3 The WDA Results

6.2.3.1 Total Exchange-Correlation Energy

Shown in Table 6.1 is the WDA total energy per electron $E_{\rm XC}^{\rm WDA}/N$, along with the difference relative to the LDA, $\Delta E_{\rm XC}^{\rm LDA}/N = (E_{\rm XC}^{\rm LDA} - E_{\rm XC}^{\rm WDA})/N$ and the GGA, $\Delta E_{\rm XC}^{\rm GGA}/N = (E_{\rm XC}^{\rm GGA} - E_{\rm XC}^{\rm WDA})/N$, for the three systems. The VMC values quoted in Ref. [161] are also given in brackets for comparison. For both the LDA and the GGA, the deviations are positive for the $q = 1.12k_F^0$ system and as the wavevector increases, $\Delta E_{\rm XC}^{\rm LDA}/N$ and $\Delta E_{\rm XC}^{\rm GGA}/N$ become negative, with the greatest differences occuring for the most inhomogeneous system. It is interesting to note that the LDA provides closer agreement with the WDA than the GGA for the $q = 1.56k_F^0$ and $q = 2.18k_F^0$ systems.

Table 6.1: The WDA total energy $E_{\rm XC}^{\rm WDA}/N$, and the percentage differences relative to the LDA and the GGA, for the three cosine-wave systems. The values in brackets are the VMC values quoted from Table 1. of Ref. [161].

q/k_F^0	$E_{\rm XC}^{\rm WDA}/N$ ($E_{\rm XC}^{\rm VMC}/N$)	$\Delta E_{\rm XC}^{\rm LDA}/N$	$\Delta E_{\rm XC}^{\rm GGA}/N$
1.12 (1.11)	−0.3327 (−0.3289)	+1.47% (+1.28%)	+0.30% (+0.03%)
1.56 (1.55)	−0.3134 (−0.3127)	−0.77% (−0.16%)	−2.46% (−2.37%)
2.18 (2.17)	−0.2874 (−0.2882)	−3.34% (−2.29%)	−5.57% (−4.86%)

The agreement between the WDA and the VMC results cannot be rigorously quantified since the densities are slightly different and also because different LDA and GGA forms where used in the VMC calculations. Nevertheless the agreement between the WDA and the VMC is certainly promising. Exactly the same trends are observed for $\Delta E_{\rm XC}^{\rm LDA}/N$ and $\Delta E_{\rm XC}^{\rm GGA}/N$ in all three systems. The mean deviations obtained from the VMC method are all greater than those of the WDA, this may be caused by the values of q/k_F^0 being slightly greater in the VMC study.

6.2.3.2 Exchange-Correlation Energy Densities

To be consistent with the VMC study the WDA energy density $e_{\rm XC}^{\rm WDA}(\mathbf{r})$ is calculated as,

$$e_{\rm XC}^{\rm WDA}(\mathbf{r}) = n(\mathbf{r})\,\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r}) = \frac{n(\mathbf{r})}{2}\int \frac{n_{\rm XC}^{\rm WDA}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\,d\mathbf{r}'\,, \tag{6.8}$$

similarly, the LDA energy density is determined as,

$$e_{\rm XC}^{\rm LDA}(\mathbf{r}) = n(\mathbf{r})\,\varepsilon_{\rm XC}^{\rm hom}(\mathbf{r})\,. \tag{6.9}$$

GGA energy densities are not presented since they cannot be rigorously defined - it is possible to add a function to $e_{\rm XC}^{\rm GGA}(\mathbf{r}) = n(\mathbf{r})\,f_{\rm XC}[n(\mathbf{r}),|\nabla n(\mathbf{r})|]$, which although leaves $E_{\rm XC}^{\rm GGA}[n(\mathbf{r})]$ unchanged, modifies $e_{\rm XC}^{\rm GGA}(\mathbf{r})$ point-wise. It should be noted that attempts have been made to correct this deficiency of the GGA [164], however, determining differences with respect to just the LDA, $\Delta e_{\rm XC}(\mathbf{r}) = e_{\rm XC}^{\rm LDA}(\mathbf{r}) - e_{\rm XC}^{\rm WDA}(\mathbf{r})$, is sufficient for the purposes of comparison with the VMC work.

Fig. 6.3 shows plots of $\Delta e_{\rm XC}$, $n(\mathbf{r})$ and the corresponding Laplacian $\nabla^2 n(\mathbf{r})$, along the direction of inhomogeneity, y, for all three systems. It is clear that the resemblance between $\Delta e_{\rm XC}(y)$ and $\nabla^2 n(\mathbf{r})$ is again observed - in extremely close agreement with the VMC results shown in Fig. 6.1. Despite the differences in the densities used in the present study, the magnitude of the deviations are also in very good agreement with the VMC results.

Fig. 6.4 displays $\Delta e_{\rm XC}(y)$ calculated for the $q = 1.12\,k_{\rm F}^0$ system using all twelve pair-correlation models described in Chapter 5, labelled WDA1 to WDA12. Except for the less physical models given by WDA7 and WDA10, all the functions exhibit the characteristic Laplacian-type deviations, but with varying amplitudes. A general trend is observed for the size of the minima that occur around the peaks in the density, in that they become larger (more negative) for the three models within each group. Other general trends are exhibited that are in complete accordance with the self-consistent properties calculated in Chapter 5, for example the smallest and largest variations in

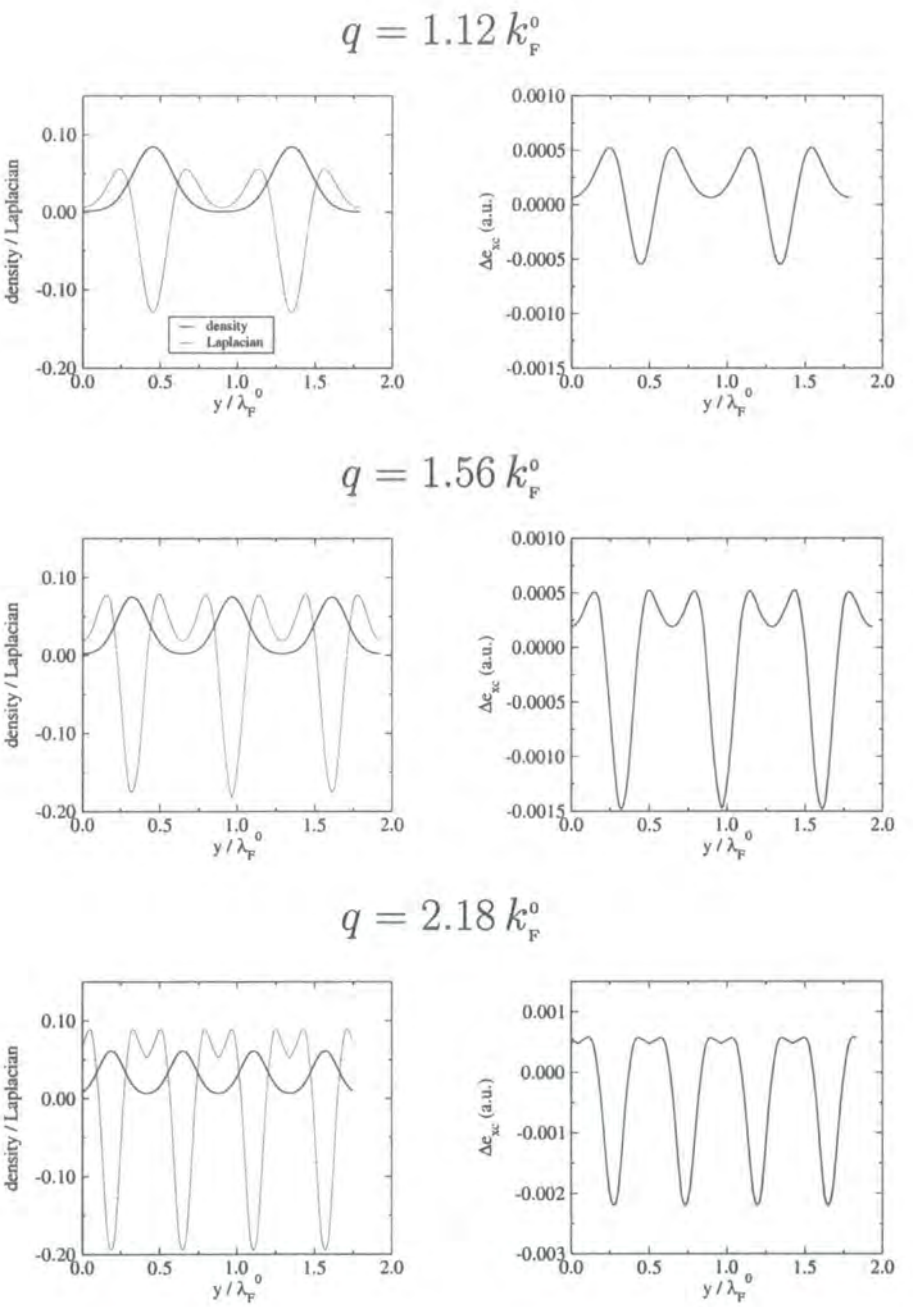

Figure 6.3: The left panels display the density together with the Laplacian of the density $\nabla^2 n(\mathbf{r})$, and the right panels show the exchange-correlation energy difference $\Delta e_{\rm XC}(y)$ = $e_{\rm XC}^{\rm LDA}(y)$ - $e_{\rm XC}^{\rm WDA}(y)$ for the cosine-wave electron gas systems with $q = 1.12\,k_{\rm F}^0$ (top), $1.56\,k_{\rm F}^0$ (middle) and $2.18\,k_{\rm F}^0$ (bottom).

$$q = 1.12\, k_F^0$$

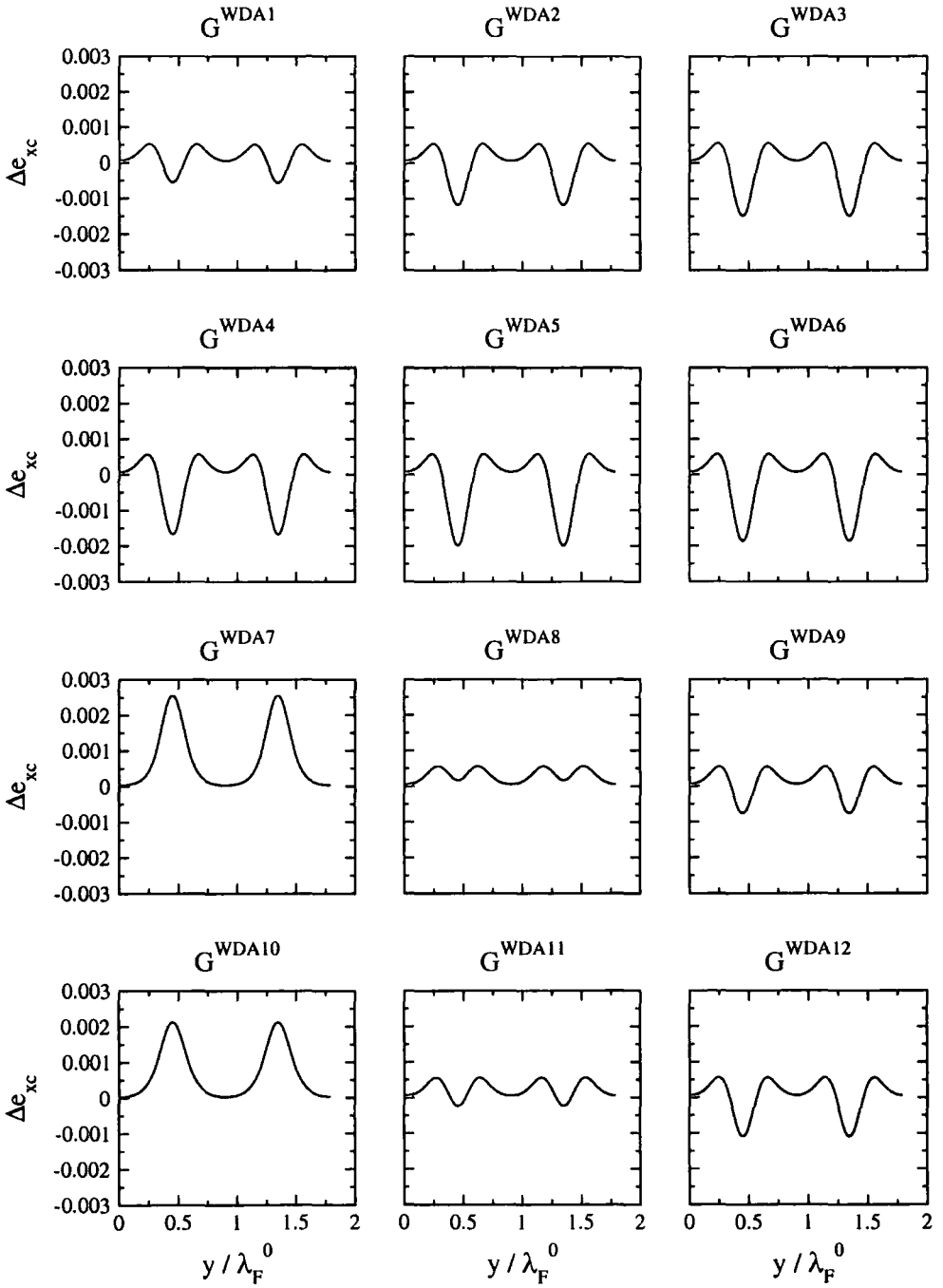

Figure 6.4: The energy density difference $\Delta e_{\rm XC}(y)$ = $e_{\rm XC}^{\rm LDA}$(y) - $e_{\rm XC}^{\rm WDA}(y)$ calculated for the $q = 1.12\, k_F^0$ system using the twelve model pair-correlation functions defined in Table 5.1.

$\Delta e_{\rm XC}(y)$ within a particular group, occur for the functions in groups 2 and 3 respectively.

6.2.3.3 Exchange-Correlation Holes

Local XC holes $n_{\rm XC}(\mathbf{r}, \mathbf{r}')$ are now examined for a reference electron located at various positions within a plane parallel to the direction of density inhomogeneity. Again, since an explicit local hole can not be obtained using the GGA, only the LDA, WDA and VMC methods can be considered. Fig. 6.5 illustrates the positions along the profile of the $q = 1.12k_F^0$ system where the holes are plotted - at the peak of a density maximum, near to a density minimum, and at a density minimum.

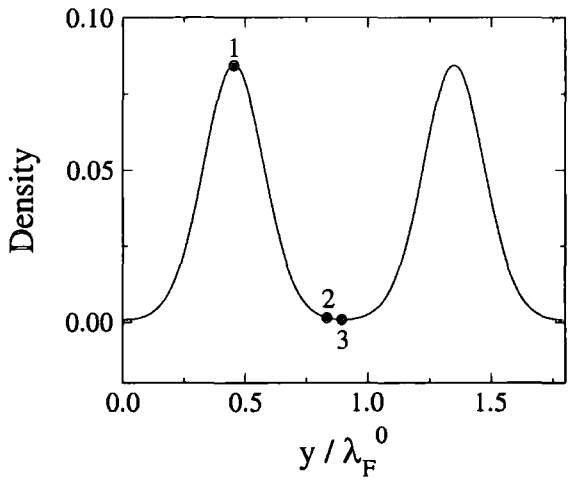

Figure 6.5: Diagram showing the positions of the electron along the $q = 1.12\,k_{\rm F}^0$ density profile where the XC holes in Fig. 6.6 are calculated. The locations are shown by filled circles and are labelled 1 (density maximum), 2 (near a density minimum) and 3 (density minimum).

When the electron is at a density maximum, shown in Fig. 6.6(a), the WDA hole is centred directly at the site of the electron and is contracted in the direction of inhomogeneity. The LDA is in good agreement with the WDA - the on-top value $n_{\rm XC}(\mathbf{r}, \mathbf{r})$ of both holes are almost identical, however

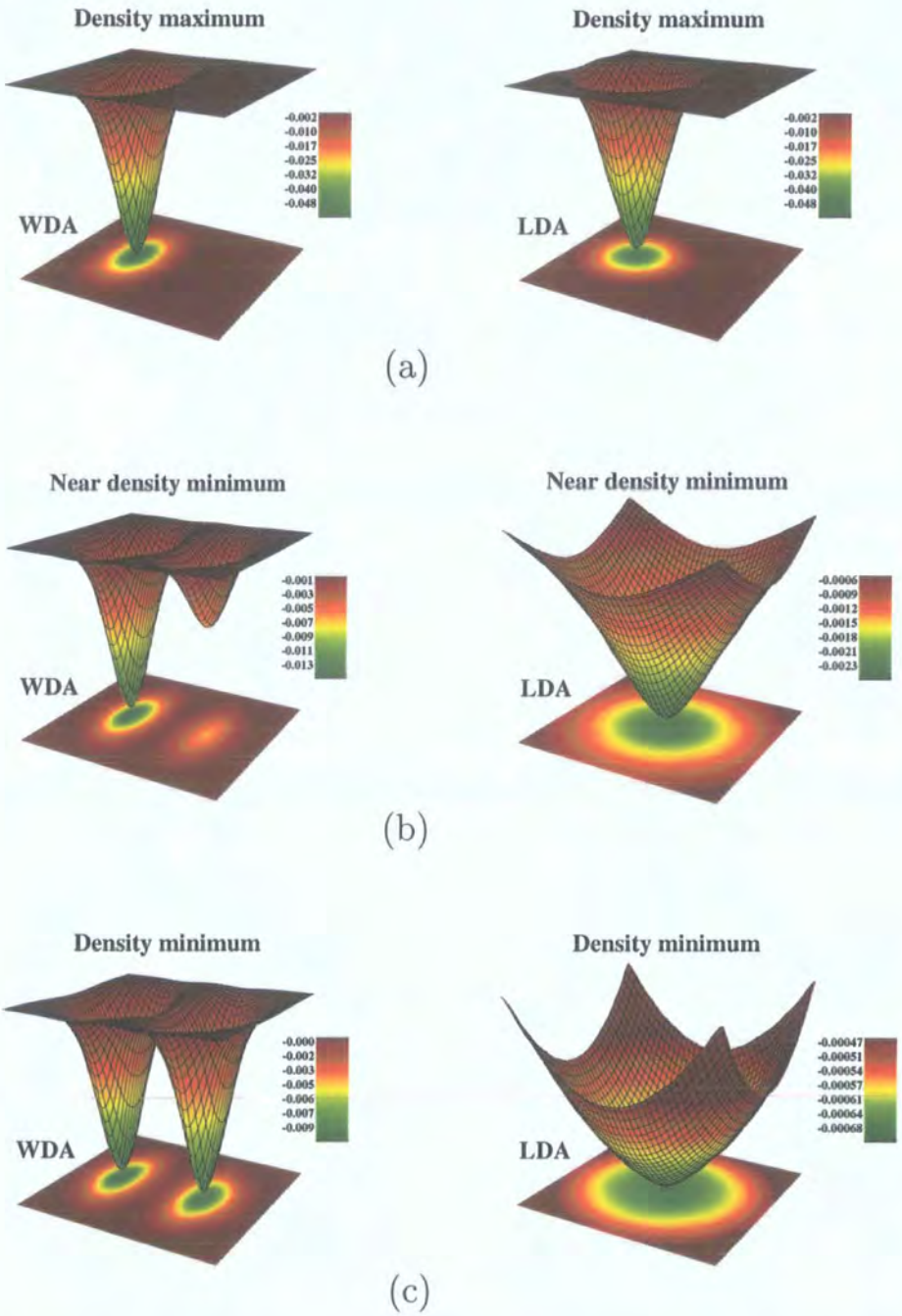

Figure 6.6: WDA and LDA exchange-correlation holes, $n_{\rm XC}(\mathbf{r}, \mathbf{r}')$, calculated when a reference electron is located at (a) a density maximum, (b) near a minimum and (c) at a density minimum (c), in the $q = 1.12\, k_{\rm F}^0$ system. The precise locations of the electron along the density profile are shown in Fig. 6.5.

as ever, the LDA hole is spherically symmetric. At positions away from the density maximum, the WDA hole is completely separated from the electron, as demonstrated in Fig. 6.6(b). In this case the hole stays fixed at the density maximum and trails behind the reference electron which is now located near the density minimum. The LDA description is now radically different from the WDA, since not only is it centred on the electron, it is considerably shallower and more diffuse. This is a general feature of LDA holes for small local density values, and is a direct consequence of satisfying the sum rule in a purely local manner. At the density minimum the hole-electron delocalisation effects are most prominent, since the WDA hole develops two minima either side of the reference electron, as shown in Fig. 6.6(c). The electron is located in the centre of the plane, equidistant from the hole minima, which are now highly contracted in the direction of inhomogeneity. In complete contrast, the LDA hole expands throughout space, extending into neighbouring unit cells.

The agreement between the WDA and VMC holes is now considered. In general, the WDA compares favourably with the VMC method - the contraction of the VMC hole along the direction of inhomogeneity reported by Nekovee *et al.* for the VMC results are also exhibited by the WDA, and the same characteristic non-local minima are observed when an electron is located at the density minimum. The depths of the VMC holes are generally greater than the WDA using the Gaussian pair-correlation function, although the difference in the densities employed may be accountable for this.

The same general features in the holes for the other two systems are observed, as illustrated in Fig. 6.7 which shows the WDA holes for an electron at a density maximum (a) and minimum (b) for the $q = 2.18\,k_{\mathrm{F}}^{0}$ system. Although, when the electron is at the peak in (a), the hole displays additional smaller minima located around at the two neighbouring peaks in the density. Another noticeable feature is that the depth of the non-local hole minima in (b) are much greater than observed at the corresponding minima in the

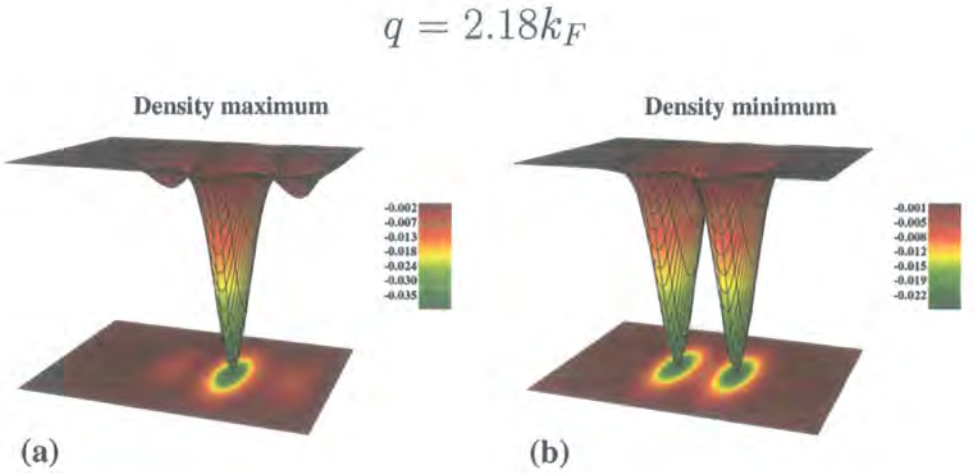

Figure 6.7: WDA exchange-correlation holes $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ calculated for a reference electron located at a density maximum, and a minimum, in the $q = 2.18\, k_{\mathrm{F}}^0$ cosine-wave system.

$q = 1.12\, k_{\mathrm{F}}^0$ system. This may account for the larger discrepancy in the total energy differences between the WDA and the LDA/GGA for this system - the value of the local density at these points is similar for both systems and so the LDA/GGA holes will remain relatively unchanged, whereas the WDA result is much different.

6.3 Quasi Two-Dimensional Electron Gas

The work in the previous section examined the effect of varying the size of the wavevector q in the cosine potential. This section explores the approach to the quasi two-dimensional (2D) electron gas limit by varying the amplitude, v_{q}, of the cosine potential which admits just one cycle of the perturbation along a single dimension of the unit cell. The behaviour of density functional approximations in the quasi-2D limit has been investigated previously by several workers [165, 166].

6.3.1 The Model Potential

The electron gas densities are again generated self-consistently from the LDA, using a model potential of the form,

$$v_{\text{ext}}(\mathbf{r}) = v_0 \cos\left(\frac{2\pi x}{l}\right) , \tag{6.10}$$

where l is the unit cell length. This yields densities with a single maximum along one of the unit cell directions. The inhomogeneity of the system, which will be characterised by the full width at half maximum (FWHM) of the density distribution, in relation to λ_{F}^0, is therefore determined by the size of the amplitude v_0. A large value of v_0 on the scale of the Fermi energy ε_{F}^0, gives rise to a narrow density profile and therefore a small FWHM. The quasi-2D limit is approached by increasing v_0 from small values up to some maximum, v_0^{max}, which results in an electron gas that is extremely confined along one direction. The value of v_0^{max} is determined when further (non-negligible) increases yield negligible changes in the self-consistent density profile.

6.3.2 Computational Results

Two systems are considered corresponding to an average density of $r_s = 2.0a_0$ and $r_s = 4.3a_0$. Examples of the density distributions are shown in Fig. 6.8.

6.3.2.1 Total Energy Differences ΔE_{XC}

Fig. 6.9 shows the differences in the total XC energy per electron relative to the WDA for the LDA ($\Delta E_{\text{XC}}^{\text{LDA}}/N$) and the GGA ($\Delta E_{\text{XC}}^{\text{GGA}}/N$) for a range of density profiles that approach the quasi-2D regime. In the $r_s = 2.0a_0$ system, shown in Fig. 6.9(a), the LDA and GGA deviations are positive for small and intermediate confinements, with the GGA being in closer agreement with the WDA. As the confinement gets stronger, the GGA differences are almost constant, whereas for the LDA they steadily increase, reaching a maximum

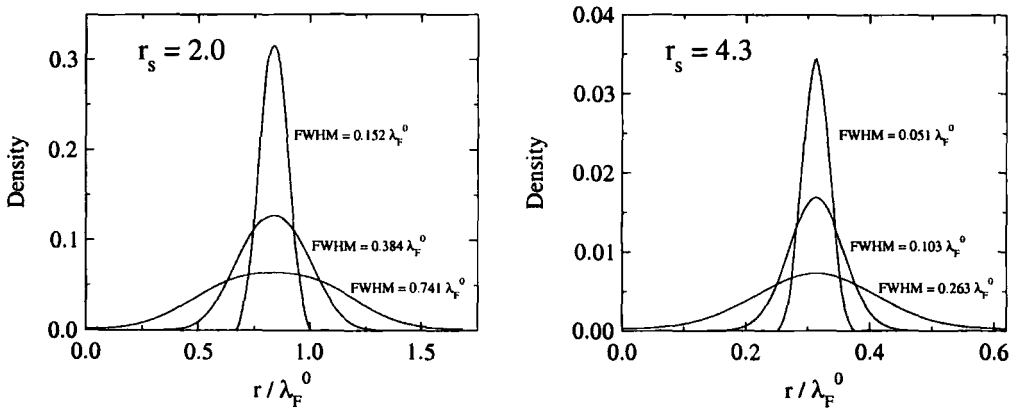

Figure 6.8: Examples of the non-uniform densities approaching the quasi-2D limit for (a) the $r_s = 2.0a_0$ system, and (b) the $r_s = 4.3a_0$ system. The distance along the direction of inhomogeneity, r, is given in terms of the Fermi wavelength λ_F^0.

when FWHM $\sim 0.25\lambda_F^0$. For stronger confinement, the differences become negative and start to diverge for both functionals, however the GGA diverges faster than the LDA. The same divergent behaviour is observed in the low density $r_s = 4.3a_0$ case given in Fig. 6.9(b), except the densities are more strongly confined on the scale of λ_F^0 in the quasi-2D regime, so the energy differences are much greater than in the high density case. These results are consistent with the findings of Kim *et al.* [165] who found that the non-local ADA gives very accurate energies in the strong 2D limit, whereas (semi)local functionals diverge to minus infinity.

6.3.2.2 Energy density Differences $\Delta e_{\rm XC}(\mathbf{r})$

Energy density differences $\Delta e_{\rm XC}(\mathbf{r})$ between the LDA and the WDA plotted along the direction of inhomogeneity are now considered. For weak confinement, $\Delta e_{\rm XC}(\mathbf{r})$ bears little resemblance to the Laplacian of the density $\nabla^2 n(r)$ as shown in Fig. 6.10(a), however for intermediately confined densities, they are strikingly similar, as shown in Fig. 6.10(b). When the density is strongly confined, $\Delta e_{\rm XC}(\mathbf{r})$ becomes large and negative near the density

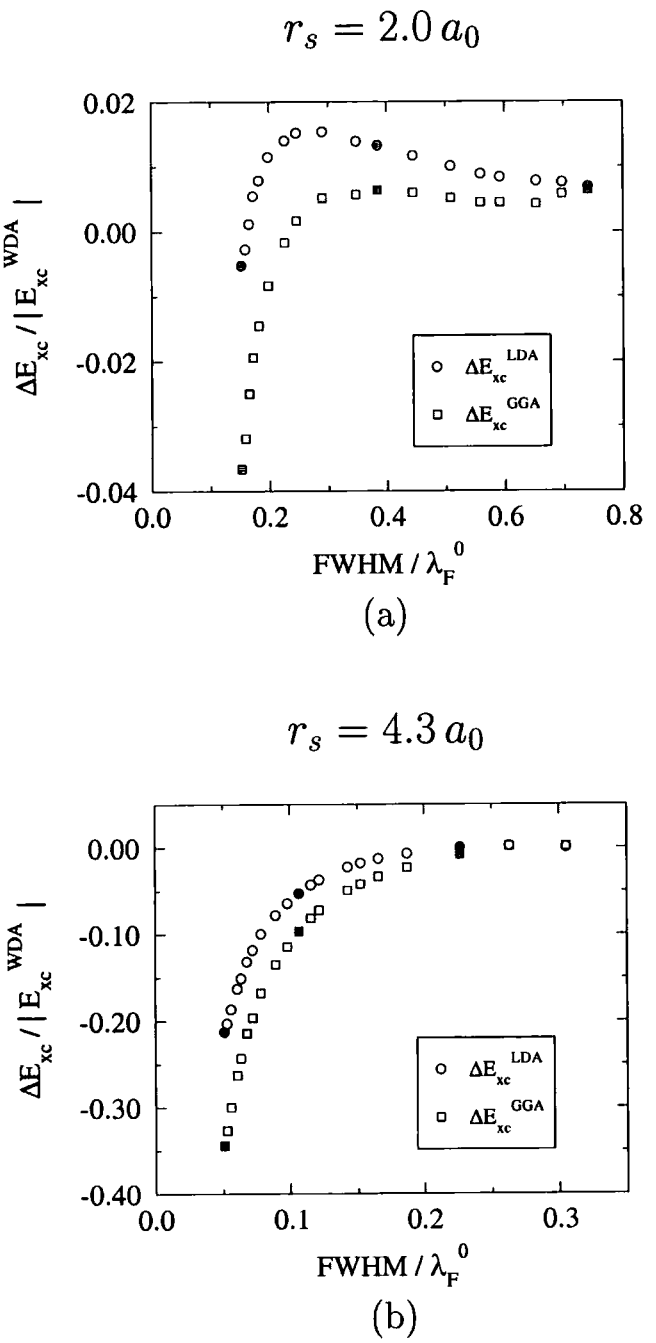

Figure 6.9: Total energy differences, $\Delta E_{\rm XC}/N$, as a function of FWHM for the LDA (circles) and GGA (boxes) relative to the WDA, for densities at (a) $r_s = 2.0a_0$ and (b) $r_s = 4.3a_0$. The shaded data points correspond to the densities shown in Fig. 6.8.

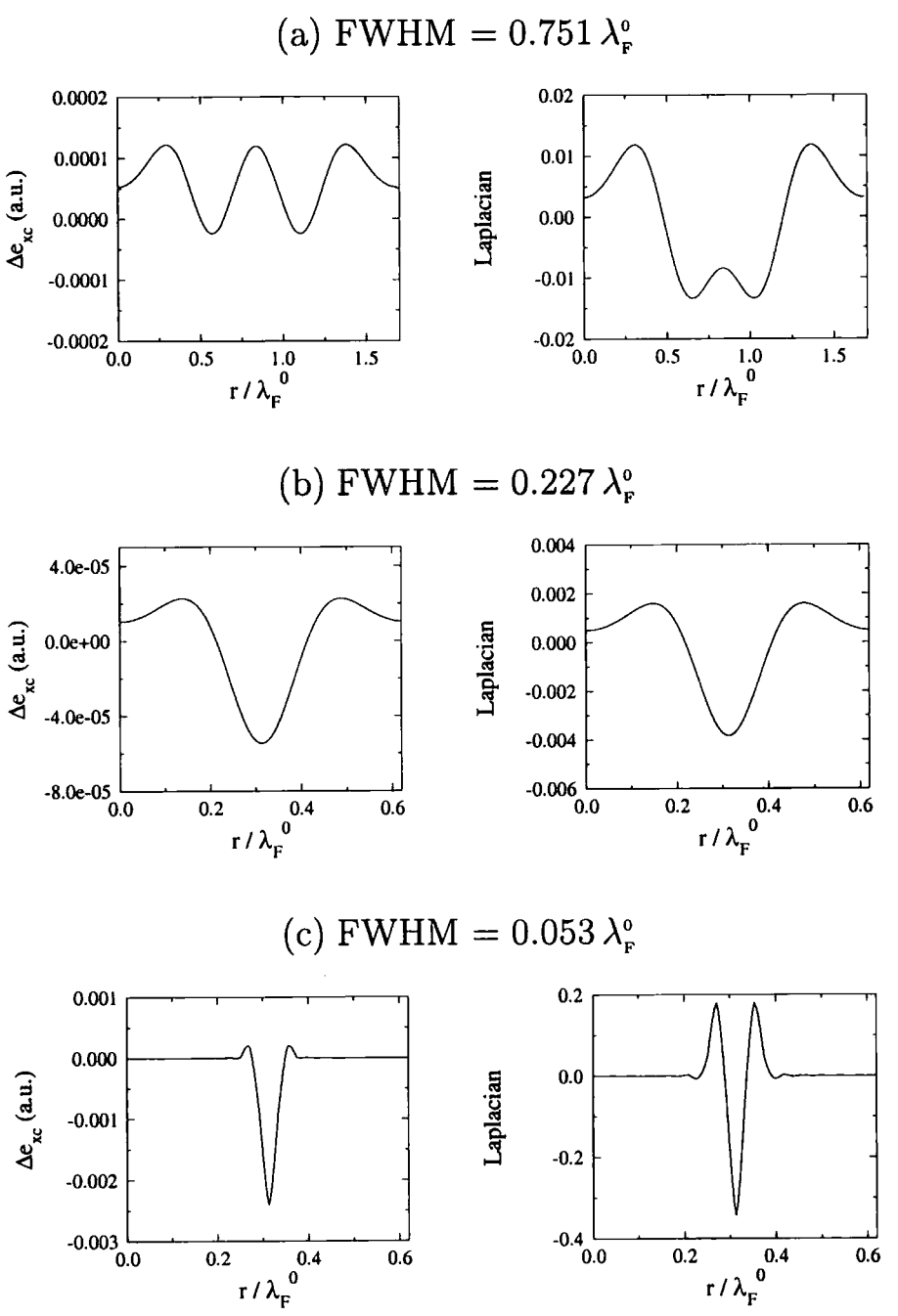

Figure 6.10: The energy density difference, $\Delta e_{\rm XC}(\mathbf{r})$, between the LDA and the WDA (left panels), and the Laplacian of the density $\nabla^2 n(\mathbf{r})$ (right panels), for electron density profiles approaching the quasi-2D electron gas limit, with (a) FWHM = $0.741\lambda_F^0$, (b) FWHM = $0.227\lambda_F^0$ and (c) FWHM = $0.053\lambda_F^0$.

maximum in comparison to $\nabla^2 n(r)$, which results in a divergence when the density is in the quasi-2D regime. The approach to this divergence is shown in Fig. 6.10(c). This last result is in accordance with the work of García-González [167] who showed that $e_{\mathrm{XC}}^{\mathrm{LDA}} \rightarrow -\infty$, when the dimensionality of an electron gas changes from 3D to 2D.

6.3.2.3 Potentials $v_{\mathrm{XC}}(\mathbf{r})$

Fig. 6.11 shows the LDA, GGA and WDA exchange-correlation potentials for the two systems in the quasi-2D regime. The LDA and GGA potentials are very similar, whereas the WDA exhibits large differences, mainly due to its slower $-1/2r$ decay. The difference is more pronounced in the $r_s = 4.3a_0$ system where the WDA potential is very shallow in comparison to the (semi)-local functionals. The effect of the WDA potential may have important consequences for the description of sub-band energy levels in real quasi-2D systems such as the inversion and accumulation layers in metal-oxide-semiconductors [168], which from a modelling viewpoint are similar to the quasi-2D electron gas densities examined here. The LDA is known to overestimate these levels in comparison to experiment [169, 170], and a shallower potential like that of the WDA is likely to provide an improvement. The LDA and GGA also exhibit large fluctuations in the tail regions of the density profiles, especially in the $r_s = 2.0a_0$ system, which are not displayed by the WDA. Spurious oscillations are known to occur with GGA potentials in instances where the dimensionless density gradient s is small [171] or large [171, 172], however since the oscillations also occur for the LDA, they are likely to be caused by an incomplete convergence of the total energy for such small densities.

6.3.2.4 Exchange-Correlation Holes $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$

Finally, XC holes are compared between the LDA and WDA for a reference electron situated at the density maximum in each system. When the

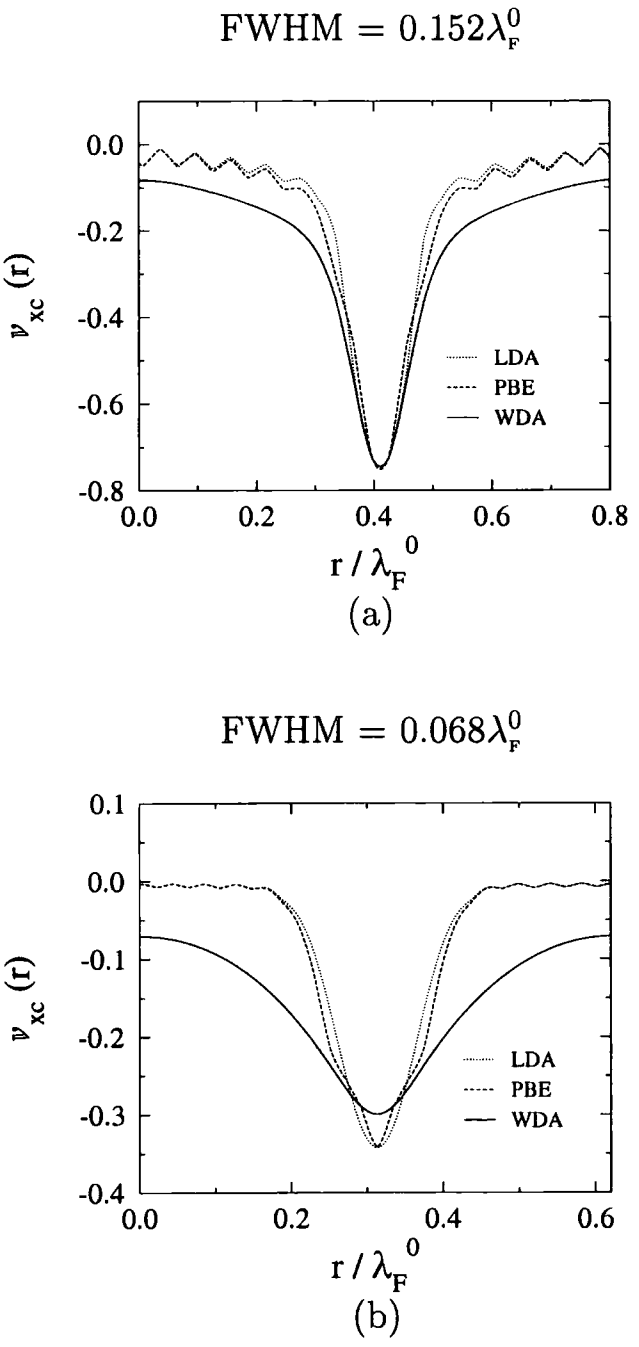

Figure 6.11: Exchange-correlation potentials $v_{\rm XC}(r)$ plotted along the direction of inhomogeneity for electron gas systems in the quasi-2D regime, (a) $r_s = 2.0$ (FWHM $= 0.152\lambda_{\rm F}^0$) and (b) $r_s = 4.3$ (FWHM $= 0.068\lambda_{\rm F}^0$).

density is sufficiently slowly varying the WDA hole is almost exactly spherical, and the LDA hole is in very good agreement as shown in Fig. 6.12 (a). However as the confinement gets stronger, the WDA hole becomes increasingly anisotropic as it contracts in the direction of the density inhomogeneity. Since the LDA hole depends on the local density, which in this case is the value at the peak in the density profile, the LDA hole remains spherical and continues to get deeper relative to the WDA as the confinement and hence the local density at the maximum increases, see Fig. 6.12 (b). This accounts for the divergent nature of the LDA energy densities, since $\epsilon_{\rm XC}(\mathbf{r})$ is directly related to the on-top hole density $n_{\rm XC}(\mathbf{r},\mathbf{r})$.

The same explanation holds for the GGA. Although an explicit local hole cannot be determined within the GGA formalism, an explanation can be given in terms of the numerical GGA hole constructed by Perdew *et al.* in Ref. [157], which reproduces the conventional analytic PBE functional. In the localised density environment studied here, exchange dominates correlation within the GGA, since as as $s \to \infty$ the cutoff radius R_c for the exchange hole (which determines the spatial extent of the hole) becomes smaller (see Fig. 2 of Ref. [157]), causing the exchange hole to become deeper and more localised, whereas the correlation contribution gradually reduces to zero [157]. As a result, the GGA exchange energy becomes increasingly negative, with $F_{\rm X}^{\rm GGA}(s)$ larger than unity, which leads to lower energies than the LDA. The GGA therefore only worsens the problem by augmenting the divergent behaviour of the LDA. The source of the problem lies with the central variable in the GGA, namely s. Since s is proportional to the modulus of the density gradient, it cannot take proper account of the fact that the density is only highly localised within one dimension, whereas the other two dimensions are completely uniform.

In contrast to the LDA and GGA, the non-local dependence on the density in the WDA allows its XC hole to distort and spread along the ridge of the density profile, resulting in shallower holes and energy densities that tend

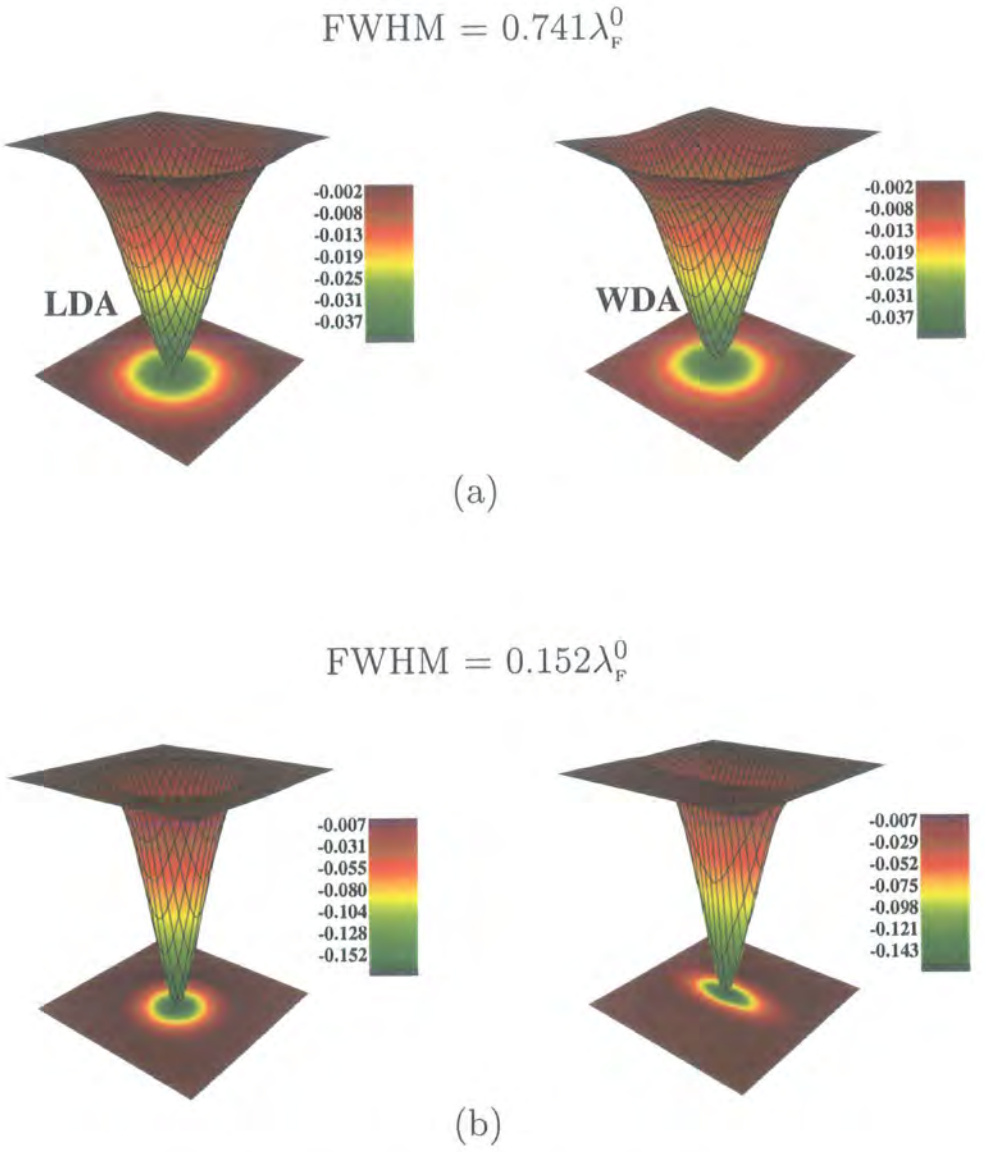

Figure 6.12: The LDA and WDA exchange-correlation hole, $n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$, surrounding an electron located at the density maximum in the $r_s = 2.0a_0$ system with (a) FWHM $= 0.741\lambda_{\mathrm{F}}^0$ and (b) FWHM $= 0.152\lambda_{\mathrm{F}}^0$.

to finite values in the strongly confined limit. A striking example of this anisotropy is illustrated in Fig. 6.13(a) for the most strongly confined density (on the scale of λ_F^0) in this study. It is clear that the WDA hole correctly takes on the same quasi-2D character as the actual density. The WDA hole can also be highly delocalised from the reference electron in this system. When the electron moves parallel to the direction of inhomogeneity, into the low density region midway between density peaks, the hole stays located at the density maximum completely delocalised from the electron. This is shown in Fig. 6.13(b).

6.3.2.5 Discussion

It is apparent that the GGA enhancement factor must be of the opposite sign for large values of s, in order to counteract the divergent behaviour in the energy density. An attempt was made in this study to achieve this using various tanh functions that become effective for large s, however it was not possible to collectively improve the total XC energy relative to the WDA for the full range of perturbations shown in Figs. 6.9(a) and 6.9(b). It appears that the inclusion of further semi-local information in the MGGA form does not resolve the problem [166], in fact the MGGA functional can be actually worse than the GGA [165].

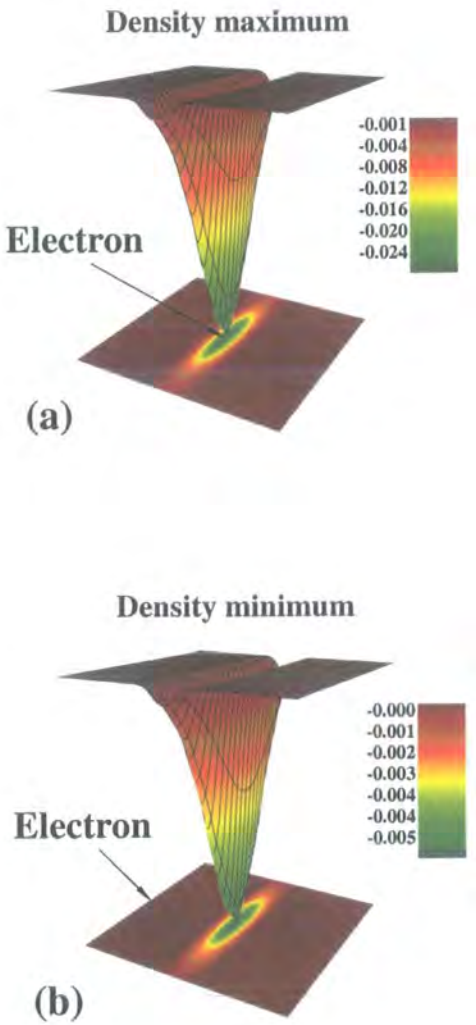

Figure 6.13: Exchange-correlation holes calculated using the WDA for a density in the quasi-2D regime (FWHM= $0.051\lambda_F^0$), when the electron is located at (a) the density maximum, and (b) at the edge of the unit cell at a density minimum. Note the different scales used in each figure. In the latter case, the hole remains fixed at the position of the density maximum, completely delocalised from the electron which is marked by the arrow.

6.4 Strong Isotropic Confinement

6.4.1 The Model Potential

As a final investigation into the properties of the WDA, an electron gas that is confined in all three dimensions with equal strength is considered. In the same manner as before, a simple cosine-potential $v_{\text{ext}}(\mathbf{r})$ is used to confine the density, except that it is applied in all three directions x, y, z of the unit cell,

$$v_{\text{ext}}(\mathbf{r}) = v_0 \cos\left[\frac{2\pi}{l}(x+y+z)\right] . \tag{6.11}$$

This model potential gives rise to a spherically symmetric confined density in the centre of the cell, with the strength of the confinement determined by the amplitude v_0. The behaviour of the LDA, GGA and WDA as the confinement increases up to some maximum value, v_{max}, will be investigated.

6.4.2 Computational Results

Two systems with the same unit cell volume are considered, and are categorised as having a high average density and low average density. The number of electrons in each case is, $N = 40$ ($r_s = 1.5a_0$), and $N = 2$ ($r_s = 4.3a_0$) respectively. The WDA calculations were performed using a $30 \times 30 \times 30$ spatial grid and 100 interpolation points for the weighted density grid.

6.4.2.1 High Average Density

To illustrate the degree of inhomogeneity encountered in this system, the density distribution determined with the strongest confining potential $v_0^{\text{max}} = 274\varepsilon_{\text{F}}^0$, is shown in Fig. 6.14. Total energy differences for the LDA and GGA relative to the WDA, $\Delta E_{\text{XC}}^{\text{LDA,GGA}}/N = (E_{\text{XC}}^{\text{LDA,GGA}} - E_{\text{XC}}^{\text{WDA}})/N$ are given in Table 6.2, for a range of v_0 values up to v_0^{max}. The deviations are small for both the LDA and the GGA for all confinement strengths - less than 1% in

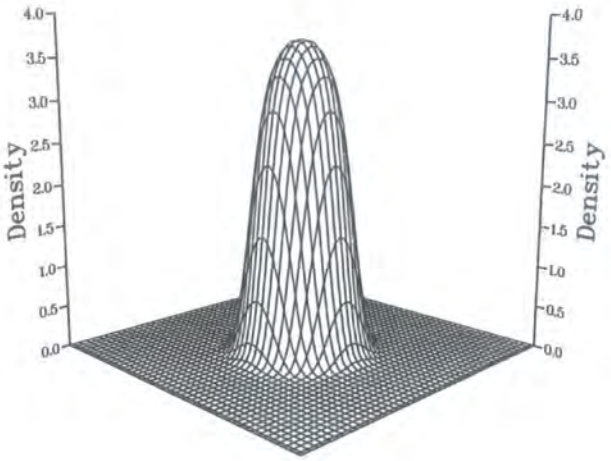

Figure 6.14: Isotropically confined electron density distribution for the system with high average density, $N = 40$ ($r_s = 1.5a_0$), plotted in a plane taken through the centre of the confining potential with $v_0 = v_0^{\max}$.

all cases. The LDA differences change sign from negative to positive as the density becomes more strongly confined, and vanishes around $v_0 = 2.7\varepsilon_{\mathrm{F}}^0$. In contrast, the GGA differences are all negative, and actually become smaller as v_0 increases. Consequently, for moderate to high values of v_0, the WDA total energy lies between the LDA and the GGA. The fact that the WDA results are generally commensurate with those of the LDA and GGA is reassuring for the WDA, since the LDA and GGA are expected to give sensible results for this type of system because the LDA and this particular GGA satisfy the exact scaling relation [56],

$$E_{\mathrm{x}}[n_\lambda] = \lambda E_{\mathrm{x}}[n]\,, \tag{6.12}$$

where n_λ represents a density that is scaled (confined) in all three dimensions and the parameter, λ, controls the amount of scaling,

$$n_\lambda(x, y, z) = \lambda^3 n(\lambda x, \lambda y, \lambda z)\,. \tag{6.13}$$

Table 6.2: The total exchange-correlation energy $E_{\rm XC}^{\rm WDA}/N$, and the difference relative to the LDA and the GGA, at several confinement strengths for the isotropically confined system with $r_s = 1.5a_0$.

$v_0/\varepsilon_{\rm F}^0$	$E_{\rm XC}^{\rm WDA}/N$	$\Delta E_{\rm XC}^{\rm LDA}/N$	$\Delta E_{\rm XC}^{\rm GGA}/N$
1.4	−0.3660	−0.0032 (−0.8%)	−0.0034 (−0.9%)
2.7	−0.4209	−0.0005 (−0.1%)	−0.0024 (−0.6%)
13	−0.6898	+0.0053 (+0.8%)	−0.0628 (−0.7%)
41	−0.9425	+0.0071 (+0.8%)	−0.0061 (−0.7%)
274	−1.0362	+0.0099 (+0.9%)	−0.0056 (−0.5%)

The system studied here is therefore an example of a uniformly scaled density since the external potential $v_{\rm ext}(\mathbf{r})$ acts on all three position coordinates. It should be noted that the exchange energy will form the dominant contribution to the total exchange-correlation energy since correlation effects diminish in the presence of strong confining potentials [165], and so satisfying the exchange energy scaling condition in (6.12) will at least lead to a good qualitative description in the strongly confined (isotropic) regime.

The exchange-correlation potential calculated with all three functionals is shown in Fig. 6.15 for the system with $v_0 = 41\varepsilon_{\rm F}^0$. As usual, the LDA and GGA display similar potentials, whereas the WDA is noticeably more slowly decaying, however it is not as extreme as in the anisotropically confined systems studied previously.

6.4.2.2 Low Average Density

In the low density case the results are very different from those just described. Fig. 6.16 shows the density distribution in a plane taken through the centre of the cell, calculated for the most strongly confined density, $v_0^{\rm max} = 2016\varepsilon_{\rm F}^0$, for this 2 electron system. Presented in Table 6.3 are the total exchange-

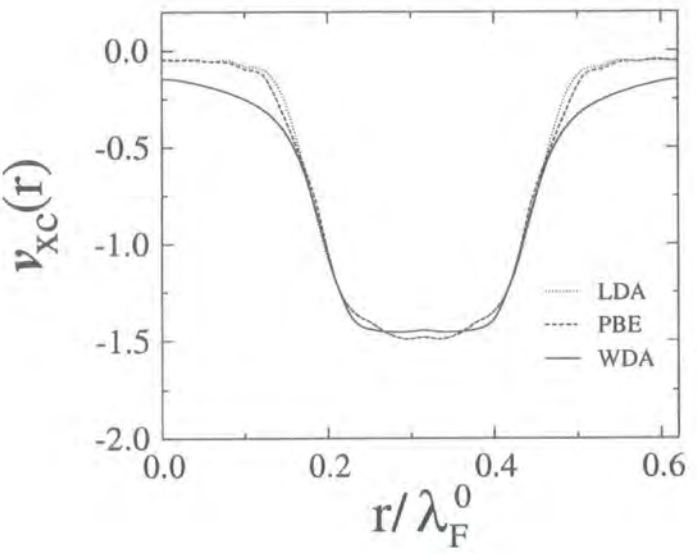

Figure 6.15: Exchange-correlation potential, $v_{\rm XC}(\mathbf{r})$, calculated with the LDA, GGA and WDA, taken through the centre of the $N = 40$ electron system, with $v_0 = 41\varepsilon_F^0$.

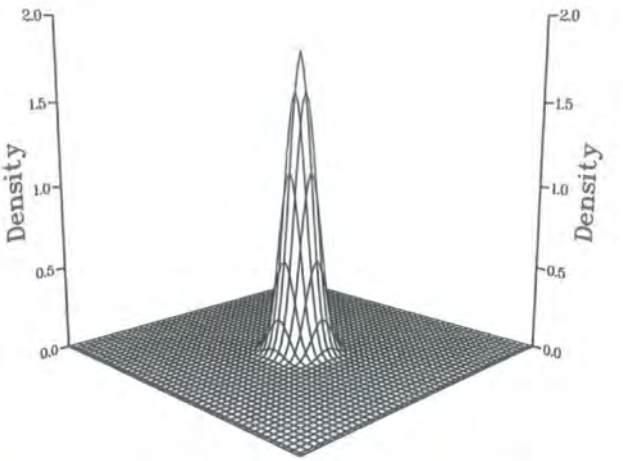

Figure 6.16: Isotropically confined electron density distribution for the system with low average density ($r_s = 4.3a_0$), plotted in a plane taken through the centre of the confining potential with $v_0 = v_0^{\max}$.

correlation energy differences $\Delta E_{\rm XC}^{\rm LDA}/N$ and $\Delta E_{\rm XC}^{\rm GGA}/N$, relative to the WDA, for several densities that span a large range of confinement strengths. The results contrast those obtained in the high density case on two counts. Firstly, the deviations are about one magnitude greater in comparison, for both functionals, although the GGA is in better agreement. Secondly, and possibly more striking, is the fact that the energy differences are hardly affected by the change in density inhomogeneity as v_0 increases. The LDA deviations vary between $10 - 12\%$, whereas for the GGA they remain almost constant at 7%, throughout the range of confinements considered.

These results can be explained by considering the self-interaction effect which is more prominent than in the high average density, simply because of the much smaller number of electrons. The WDA contains a more accurate account of self-interaction effects than the LDA and GGA, as described in Sec. 4.2.4, and if the WDA is considered to be close to the exact result, then Table 6.3 shows that the GGA provides an improvement over the LDA for self-interaction errors. Also, other than being coincidental, the distinct lack of variation in the total energy differences as the amount of density

Table 6.3: The total exchange-correlation energy $E_{\rm XC}^{\rm WDA}/N$, and the difference relative to the LDA and the GGA, for various confinement strengths in the $r_s = 4.3a_0$ system.

$v_0/\varepsilon_{\rm F}^0$	$E_{\rm XC}^{\rm WDA}/N$	$\Delta E_{\rm XC}^{\rm LDA}/N$	$\Delta E_{\rm XC}^{\rm GGA}/N$
2.0	−0.3321	+0.0336 (+10%)	+0.0230 (+7%)
20	−0.4042	+0.0452 (+11%)	+0.0305 (+8%)
101	−0.6155	+0.0740 (+11%)	+0.0455 (+7%)
303	−0.7789	+0.0935 (+12%)	+0.0534 (+7%)
2016	−0.8831	+0.8530 (+12%)	+0.0581 (+7%)

localisation increases, indicates that the self-interaction error is overwhelming the error caused by the inhomogeneity in the density.

Once more, the source of the discrepancies between the different functionals can be rationalised in terms of their respective descriptions of the exchange-correlation hole. When an electron moves out from the main density distribution in the centre of the cell, into the tail of the density, the LDA hole, as always, stays centred on the electron, whereas the WDA hole will stay localised at the density peak in the centre. A clear demonstration of this effect is given in Fig. 6.17 which shows the WDA hole for an electron at three positions moving from the centre to one of the corners of the unit cell. From the experience gained in the previous chapters, it is apparent that the electron will become separated from its hole as soon as it moves away from the central density distribution, and when the electron is situated in the corner of the cell, as in Fig. 6.17(c), its hole is located over 3Å away (since $l = 4.71$Å for this system). Also, despite the extremely low value of the local density in this region, $r_s \sim 75a_0$ at this point, the WDA hole can still be observed on the same scale as when it is situated at the density peak in Fig. 6.17(a), where the local density has a value $r_s \sim 0.7a_0$. The LDA will

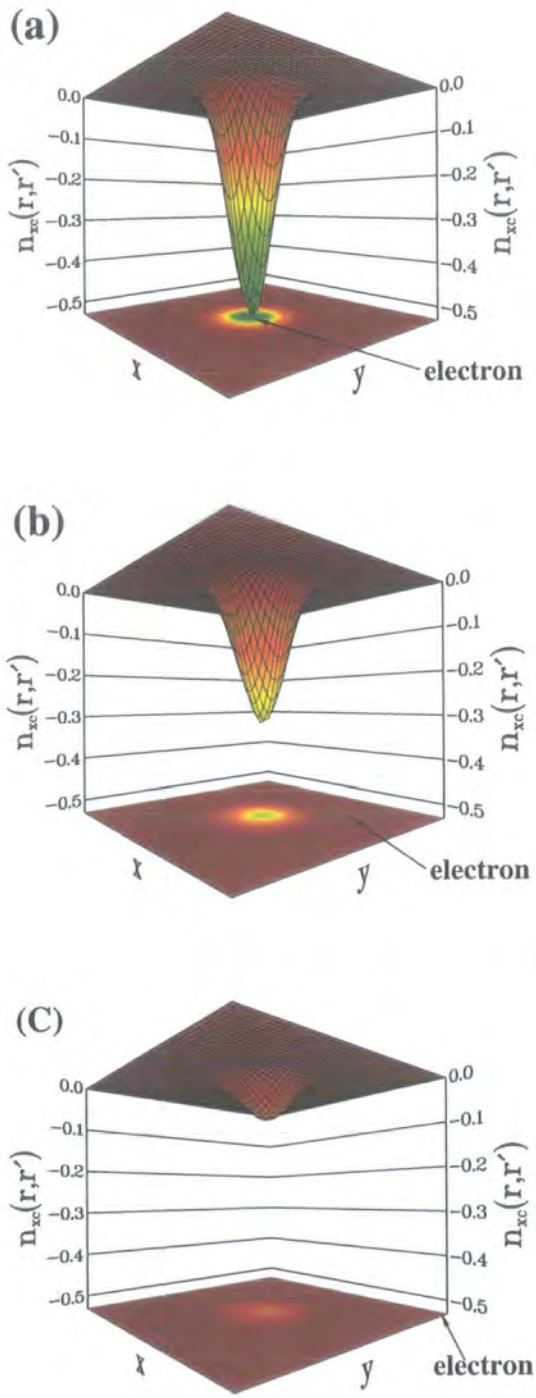

Figure 6.17: WDA exchange-correlation holes in the $N = 2$ electron system with $v_0 = v_0^{\max}$, calculated for an electron located at three points, moving along a diagonal direction from the centre of the density distribution, to a corner of the unit cell. The position of the electron in the $x - y$ plane is marked in each case.

therefore give an underestimated (less negative) value for the energy-density $\varepsilon_{\rm XC}(\mathbf{r})$ compared to the WDA, since its hole will be shallower than the WDA as a result of sampling only the local density $n(\mathbf{r})$. The importance of retaining the non-local density dependence $n(\mathbf{r}')$ in the formulation of a model for the XC hole cannot be overstated in these circumstances.

As a note, it may appear that the hole in Fig. 6.17(c) does not satisfy the sum rule when compared with the holes presented in Figs. 6.17(a) and (b). The reason is because the electron is situated at the corner of the unit cell, so from periodic boundary conditions there are hole contributions emanating from the neighbouring unit cells that are not present when the electron is located near the centre of the cell.

Fig. 6.18 shows the energy-density difference $\Delta\varepsilon_{\rm XC}(\mathbf{r}) = \varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r}) - \varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ calculated in a plane going through the centre of the cell. The fact that $\Delta\varepsilon_{\rm XC}(\mathbf{r})$ is positive at all points in the plane demonstrates that $\varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ is indeed more negative than $\varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r})$, for the reasons just given concerning the XC holes. The positive total energy differences given in Table 6.3 are therefore explained. It is presumed that the GGA is behaving in a similar way as the LDA, although the GGA description of the hole appears to be marginally better, judging from the closer agreement with the WDA total energies in Table 6.3.

6.5 Summary

The inhomogeneous electron gas was chosen as the basis of this chapter because it allows a controlled environment to examine the properties of exchange-correlation functionals, that is also free from the pseudopotential approximation. The work on the one-dimensional cosine-wave electron gas in Sec. 6.2 demonstrated the closely similarities between the WDA and the VMC method, which is highly promising. It is therefore hoped that the success of this study will stimulate further quantum Monte Carlo work so

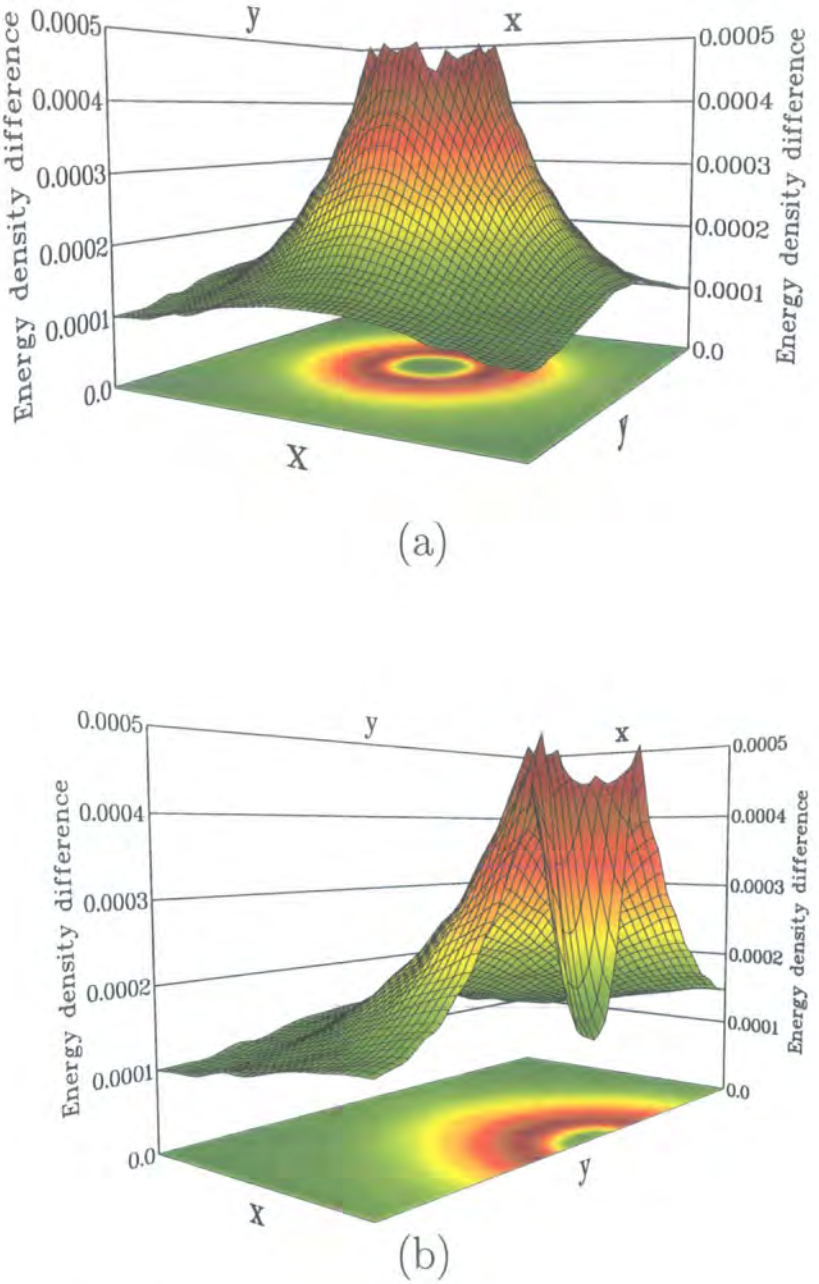

Figure 6.18: The difference between the LDA and WDA energy densities, $\Delta\varepsilon_{\rm XC}(\mathbf{r}) = \varepsilon_{\rm XC}^{\rm LDA}(\mathbf{r}) - \varepsilon_{\rm XC}^{\rm WDA}(\mathbf{r})$ (in Hartrees), calculated in a plane through the centre of the cell, for the $N = 2$ electron system with $v_0 = v_0^{\rm max}$. $\Delta\varepsilon_{\rm XC}$ is plotted throughout the entire plane in (a), whereas just half of the plane is displayed in (b), in order to expose the behaviour near the centre of the density.

that the WDA can be accurately tested in a broad variety of density environments. These comparisons can continue without the use of consistent pseudopotentials, so long as the same, or very similar, reference densities are used.

The quasi-2D system studied in Sec. 6.3 demonstrated the divergent trend of the LDA and GGA total energies and energy densities with respect to the WDA, which results from the inadequacy of a local/semi-local description of the exchange-correlation hole. This demonstrates the possibility of successfully applying the WDA to systems that exhibit low dimensional character such as semiconductor quantum well structures which have important device applications [168].

Finally, in the case of isotropic confinement examined in Sec. 6.4, the WDA yielded similar results as the LDA and GGA for the system containing a large number of electrons. However when there are few electrons in the system, the LDA and GGA underestimate total energies quite substantially in comparison to the WDA. This was discussed in the context of self-interaction effects, but is ultimately a direct consequence of the non-locality of the exchange-correlation hole, which was demonstrated by the WDA. An interesting application of the WDA would be to the determination of molecular reaction barriers and the calculation of corresponding exchange-correlation holes due to the predominance of self-interaction effects in these systems, especially since conventional functionals yield large errors for most reactions.

Chapter 7

Conclusions and Future Work

The aim of this thesis was to develop the approximation used for exchange and correlation in density functional theory. This is a fundamental problem in DFT since the very accuracy of this formally exact theory rests on this single quantity. There have been two notable advances in the history of exchange-correlation functional development. The first was the completion of a computationally viable LDA in 1981, made possible by quantum Monte Carlo simulations of the homogeneous electron gas. The second came with the introduction of the GGA in the mid 1980s, of which the most common formulations used presently have their origins in the early 1990s. The GGA has since become the accepted functional in DFT calculations within condensed matter physics.

There are many GGA functionals in existence, however the GGA formalism as a whole possesses numerous shortcomings that are well documented - these essentially emanate from the LDA on which it is an extension. Nevertheless, leaving aside such instances where the GGA fails completely, the work conducted in Chapter 3 using a highly flexible GGA functional demonstrated the difficulty in devising a semi-empirical GGA that consistently improves upon non-empirical GGA constructions, for quantities that the GGA is known to work well for, namely structural and cohesive properties of simple semiconductors. This lends credence to the view that the GGA has reached

the limits of its capabilities, and that other types of functional should now be explored.

The established strategy for superseding the GGA [173] is to incorporate further semi-local information in the form of the Laplacian of the density, and the orbital dependent kinetic energy density. Apart from problematic computational issues associated with these MGGA functionals, there still remain deep-seated theoretical deficiencies that are simply not addressed, and yet are often a serious source of errors for a variety of physical properties. The recurring problems, such as the self-interaction error and the incorrect asymptotic form of the exchange-correlation potential, are relics of a local/semi-local prescription and can only be resolved with a fully non-local functional for exchange and correlation. Also, no amount of higher order semi-local information or even a mixture of exact (Hartree-Fock) exchange can treat important, albeit weak, long-range correlated phenomena such as those caused by van der Waals interactions. At the other extreme, the description of strongly correlated materials would be equally futile with such severe approximations. Faced with mounting evidence against semi-local functionals, it would not be inappropriate to state that a concerted effort toward practical non-local functionals is an inevitability, if not long overdue. The work on the non-local weighted density approximation presented here has hopefully brought this eventuality one step closer.

Although the theory behind the WDA is very simple, its computational implementation is not so straightforward. Nevertheless an efficient algorithm for the WDA, based on a reciprocal-space representation, was developed and demonstrated in Chapter 4, which also permits new model pair-correlation functions to be incorporated with relative ease. The success of this method however should not be overstated - this implementation is still too time consuming for most practical purposes, in comparison to the cheaper alternatives such as GGA. The computational bottleneck is the determination of the weighted density throughout space, therefore the development of faster

ways to calculate $\tilde{n}(\mathbf{r})$ is one of the most important challenges facing the WDA if it is to be taken out of the realm of specialised individual users and into the broader academic and industrial arenas.

The work on the exchange-correlation holes presented in Chapter 5 for bulk Si, and in Chapter 6 for the model electron gas systems, demonstrated the valuable insight into the WDA that can be gained from this quantity, especially when compared with accurate data from variational Monte-Carlo simulations. The close agreement between the WDA and VMC results, even for strongly inhomogeneous density regimes, indicates that the spherical nature of the model pair-correlation functions does not appear to inhibit the performance of the WDA greatly. However further studies must be performed to ascertain the range of validity of this approximation. It is therefore hoped that VMC simulations performed in the future will be dedicated to testing the WDA in a broader range of systems, and also to developing the pair-correlation function approximation.

It can be easily ascertained from the work contained in this thesis that the development of the WDA is far from complete, in fact it is possibly still in its infancy. There are still very challenging theoretical problems to be tackled within the WDA, such as developing a correctly symmetrised pair-correlation function, i.e. one that possesses the property $g_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = g_{\mathrm{xc}}(\mathbf{r}',\mathbf{r})$. Also, in order to assess the performance of the WDA for real material properties, it is necessary to create consistent pseudopotentials, which may or may not involve shell partitioning techniques [85]. However, it is clear from the present state of development that the WDA will play an important role in exchange-correlation functional development in the future.

Appendix A

Derivation of the Exact Exchange-Correlation Energy

A definition of the exact exchange-correlation energy within Kohn-Sham DFT is provided by a method known as adiabatic connection, or coupling constant integration, and is described below. A detailed description of the method is given in Ref. [134].

Consider a system in which the strength of the electron-electron interaction is dependent upon a parameter λ, that effectively couples the interaction strength to the magnitude of the electron charge, $e^2 \rightarrow \lambda e^2$. The adiabatic connection procedure scales λ (and hence the electron interaction), between the real system ($\lambda = 1$) and the non-interacting case ($\lambda = 0$), whilst keeping the density $n(\mathbf{r})$ fixed (hence the reason for the term adiabatic). This is achieved using a Hamiltonian in which the electron-electron operator is multiplied by λ, and a λ-dependent potential, $\hat{V}_\lambda$, is added which maintains the fixed density constraint for all values of λ,

$$\hat{H}_\lambda = \hat{T} + \lambda \hat{V}_{ee} + \hat{V}_\lambda \tag{A.1}$$

and

$$\hat{V}_\lambda = \sum_{i=1}^{N} v_\lambda(\mathbf{r}) \,. \tag{A.2}$$

When $\lambda = 1$ the physical Hamiltonian is obtained and $\hat{V}_\lambda$ becomes identical to the external potential $\hat{V}_{\text{ext}}$. The total energy of the system at the limits of the adiabatic connection can be written simply as,

$$E_{\lambda=1}[n(\mathbf{r})] \; - \; E_{\lambda=0}[n(\mathbf{r})] = \int_0^1 dE_\lambda \,, \tag{A.3}$$

and since there is no explicit λ-dependence in the wavefunctions Ψ_λ,

$$\frac{\partial}{\partial\lambda}\langle\Psi_\lambda|\hat{H}_\lambda|\Psi_\lambda\rangle = \langle\Psi_\lambda|\partial\hat{H}_\lambda/\partial\lambda|\Psi_\lambda\rangle \,, \tag{A.4}$$

and so the total energy of the fully interacting system can be written as,

$$E_{\lambda=1}[n(\mathbf{r})] = \langle\Psi_{\lambda=0}|\hat{H}_{\lambda=0}|\Psi_{\lambda=0}\rangle + \int_0^1 \frac{d\langle\Psi_\lambda|\hat{H}_\lambda|\Psi_\lambda\rangle}{d\lambda}\, d\lambda \,, \tag{A.5}$$

where Ψ_λ is the groundstate of $\hat{H}_\lambda$. Using the Hellmann-Feynman theorem [105, 106], the integrand can be written as

$$\frac{d\langle\Psi_\lambda|\hat{H}_\lambda|\Psi_\lambda\rangle}{d\lambda} = \frac{1}{2}\int \frac{P_\lambda(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}\, d\mathbf{r}' + \int n(\mathbf{r})\frac{dv_\lambda}{d\lambda}\, d\mathbf{r} \,, \tag{A.6}$$

where the pair-density $P_\lambda(\mathbf{r},\mathbf{r}')$ is determined from Ψ_λ. The Kohn-Sham non-interacting system is also obtained for $\lambda = 0$, so that

$$\langle\Psi_{\lambda=0}|\hat{H}_{\lambda=0}|\Psi_{\lambda=0}\rangle = T_s + \int n(\mathbf{r})\, v_{\lambda=0}(\mathbf{r})\, d\mathbf{r} \,. \tag{A.7}$$

T_s is the kinetic energy of the non-interacting system determined from the wavefunction $\Psi_{\lambda=0}$, as calculated in the Kohn-Sham scheme, i.e.

$$T_s = \langle\Psi_{\lambda=0}\,|\,\hat{T}\,|\,\Psi_{\lambda=0}\,\rangle \,. \tag{A.8}$$

Substituting Eqs. (A.6) and (A.7) into (A.5) yields an expression for the true interacting system,

$$E_{\lambda=1}[n(\mathbf{r})] \;=\; T_s + \int n(\mathbf{r})\, v(\mathbf{r})\, d\mathbf{r} + \frac{1}{2}\int_0^1 d\lambda \int \frac{P_\lambda(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}\, d\mathbf{r}' \,. \tag{A.9}$$

Partitioning $P_\lambda(\mathbf{r},\mathbf{r}')$ in the manner of (2.7), with a λ-independent classical term and a λ-dependent exchange-correlation hole $n_{\text{XC},\lambda}(\mathbf{r},\mathbf{r}')$, such that,

$$P_\lambda(\mathbf{r},\mathbf{r}') = n(\mathbf{r})n(\mathbf{r}') + n(\mathbf{r})\, n_{\text{XC},\lambda}(\mathbf{r},\mathbf{r}') \,, \tag{A.10}$$

leads to an energy expression for the fully interacting system that is in the same form as the corresponding total energy expression in Kohn-Sham DFT, given by relations (1.30) and (1.45),

$$E_{\lambda=1}[n(\mathbf{r})] = T_s + \int n(\mathbf{r})\, v(\mathbf{r})\, d\mathbf{r} + \frac{1}{2}\int\int \frac{n(\mathbf{r})\, n(\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}\, d\mathbf{r}' + \int_0^1 E_{\mathrm{xc},\lambda}[n(\mathbf{r})]\, d\lambda\,, \tag{A.11}$$

where the λ-dependent exchange-correlation energy expression is given by,

$$E_{\mathrm{xc},\lambda}[n(\mathbf{r})] = \frac{1}{2}\int n(\mathbf{r})\, d\mathbf{r} \int \frac{n_{\mathrm{xc},\lambda}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}'\,. \tag{A.12}$$

So the exact exchange-correlation energy functional in DFT is defined as,

$$E_{\mathrm{xc}}[n(\mathbf{r})] = \int_0^1 E_{\mathrm{xc},\lambda}[n(\mathbf{r})]\, d\lambda = \frac{1}{2}\int n(\mathbf{r})\, d\mathbf{r} \int \frac{n_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')}{|\,\mathbf{r}-\mathbf{r}'\,|}\, d\mathbf{r}'\,, \tag{A.13}$$

where $n_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')$ is the coupling-constant averaged pair-correlation function,

$$n_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = \int_0^1 n_{\mathrm{xc},\lambda}(\mathbf{r},\mathbf{r}')\, d\lambda\,. \tag{A.14}$$

The adiabatic connection method not only provides an exact definition for $E_{\mathrm{xc}}[n(\mathbf{r})]$ through relation (A.13), it also states the significant result that the difference between the non-interacting Kohn-Sham kinetic energy T_s, and the physical kinetic energy T, is incorporated into the definition of $E_{\mathrm{xc}}[n(\mathbf{r})]$.

Appendix B

Coefficients for HCTH and HCTH-related GGAs

Table B.1: Expansion coefficients that define HCTH [98], HCTH-HEG and HCTH-26 GGA functionals.

Coefficient	HCTH	HCTH-HEG	HCTH-26
$c_{\mathrm{X}\sigma,0}$	$0.109320D+01$	$0.100000D+01$	$0.109951D+01$
$c_{\mathrm{C}\sigma\sigma,0}$	$0.222601D+00$	$0.100000D+01$	$0.468314D-01$
$c_{\mathrm{C}\alpha\beta,0}$	$0.729974D+00$	$0.100000D+01$	$0.832308D+00$
$c_{\mathrm{X}\sigma\sigma,1}$	$-0.744056D+00$	$0.835193D+00$	$-0.808922D+00$
$c_{\mathrm{C}\sigma\sigma,1}$	$-0.338622D-01$	$0.376859D+01$	$0.191901D+01$
$c_{\mathrm{C}\alpha\beta,1}$	$0.335287D+01$	$-0.388880D+01$	$0.447358D+01$
$c_{\mathrm{X}\sigma,2}$	$0.559920D+01$	$-0.185388D+01$	$0.711120D+01$
$c_{\mathrm{C}\sigma\sigma,2}$	$-0.125170D-01$	$-0.109625D+02$	$-0.329297D+01$
$c_{\mathrm{C}\alpha\beta,2}$	$-0.115430D+02$	$0.246494D+02$	$-0.234711D+02$
$c_{\mathrm{X}\sigma,3}$	$-0.678549D+01$	$0.765176D+01$	$-0.111952D+02$
$c_{\mathrm{C}\sigma\sigma,3}$	$-0.802496D+00$	$0.117238D+02$	$0.251747D+01$
$c_{\mathrm{C}\alpha\beta,3}$	$0.808564D+01$	$-0.495992D+01$	$0.340437D+02$
$c_{\mathrm{X}\sigma,4}$	$0.449357D+01$	$-0.536148D+01$	$0.901788D+01$
$c_{\mathrm{C}\sigma\sigma,4}$	$0.155396D+01$	$0.333213D+02$	$0.123873D+00$
$c_{\mathrm{C}\alpha\beta,4}$	$-0.447857D+01$	$-0.130326D+01$	$-0.227934D+02$

Table B.2: Expansion coefficients that define HCTH-$\frac{1}{4}$ [131], HCTH-26-$\frac{1}{4}$ and HCTH-HEG-26-$\frac{1}{4}$ GGA functionals.

Coefficient	HCTH-$\frac{1}{4}$	HCTH-26-$\frac{1}{4}$	HCTH-HEG-26-$\frac{1}{4}$
$c_{\mathrm{X}\sigma,0}$	$0.103161D+01$	$0.104633D+01$	$0.100000D+01$
$c_{\mathrm{C}\sigma\sigma,0}$	$0.282414D+01$	$0.185936D+01$	$0.100000D+01$
$c_{\mathrm{C}\alpha\beta,0}$	$0.821827D-01$	$0.252306D-01$	$0.100000D+01$
$c_{\mathrm{X}\sigma\sigma,1}$	$-0.360781D+00$	$-0.461495D+00$	$0.234527D+00$
$c_{\mathrm{C}\sigma\sigma,1}$	$0.318843D-01$	$-0.164315D+00$	$0.708237D+00$
$c_{\mathrm{C}\alpha\beta,1}$	$0.456466D+01$	$0.299618D+01$	$-0.797010D+00$
$c_{\mathrm{X}\sigma,2}$	$0.351994D+01$	$0.498870D+01$	$0.325534D+01$
$c_{\mathrm{C}\sigma\sigma,2}$	$-0.178512D+01$	$0.891443D+00$	$-0.103809D+01$
$c_{\mathrm{C}\alpha\beta,2}$	$-0.135529D+02$	$-0.175184D+02$	$-0.882719D+01$
$c_{\mathrm{X}\sigma,3}$	$-0.495944D+01$	$-0.759118D+01$	$-0.516729D+01$
$c_{\mathrm{C}\sigma\sigma,3}$	$0.239795D+01$	$-0.268928D+01$	$0.575599D-01$
$c_{\mathrm{C}\alpha\beta,3}$	$0.133820D+02$	$0.251991D+02$	$0.147845D+02$
$c_{\mathrm{X}\sigma,4}$	$0.241165D+01$	$0.463102D+01$	$0.341337D+01$
$c_{\mathrm{C}\sigma\sigma,4}$	$-0.876909D+00$	$0.276490D+01$	$0.991168D+00$
$c_{\mathrm{C}\alpha\beta,4}$	$-0.317493D+01$	$-0.126500D+02$	$-0.811301D+01$

Appendix C

Publications

The following is list of papers that have been published, or submitted for publication, as a result of the research carried out in this thesis.

- *Density functional calculations of semiconductor properties using a semi-empirical exchange-correlation functional*
  Phys. Rev. B **63**, 115206 (2001)
  Philip P. Rushton, Stewart J. Clark and David J. Tozer

- *Description of exchange and correlation in the strongly inhomogeneous electron gas using a non-local density functional*
  Phys. Rev. B **65**, 193106 (2002)
  Philip P. Rushton, David J. Tozer and Stewart J. Clark

- *Non-local density functional description of exchange and correlation in silicon*
  Phys. Rev. B **65**, 235202 (2002)
  Philip P. Rushton, David J. Tozer and Stewart J. Clark

- *Comparison of exchange-correlation functionals in the strongly confined electron gas*
  Submitted to Phys. Rev. B
  Philip P. Rushton and Stewart J. Clark

- *Demonstrating the effectiveness of a non-local density functional description of exchange and correlation*
  To appear in, Recent Progress in Computational Chemistry and Physics
  Philip P. Rushton and Stewart J. Clark

Bibliography

[1] E. Schrödinger, Ann. Physik **79**, 361 (1926).

[2] M. Born and R. Oppenheimer, Ann. Phys. **84**, 457 (1927).

[3] L. H. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927).

[4] E. Fermi, Z. Phys. **48**, 73 (1928).

[5] V. Fock, Z. Phys. **61**, 126 (1930).

[6] J. C. Slater, Phys. Rev. **81**, 385 (1951).

[7] D. R. Hartree, Proc. R. Soc. London **A113**, 621 (1928).

[8] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864B (1964).

[9] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133A (1965).

[10] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[11] G. D. Mahan, *Many-Particle Physics* (Plenum Press, New York and London, 1990).

[12] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[13] E. Teller, Rev. Mod. Phys. **34**, 627 (1962).

[14] N. Balázs, Phys. Rev. **156**, 42 (1967).

[15] E. H. Lieb and F. Y. Wu, Phys. Rev. Lett. **31**, 681 (1968).

[16] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).

[17] P.-O. Löwdin, Adv. Chem. Phys. **2**, 207 (1959).

[18] R. J. Bartlett and J. F. Stanton, Rev. Comput. Chem. **5**, 65 (1995).

[19] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, 1986).

[20] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Harcourt College Publishers, 1976).

[21] A. Balereschi, Phys. Rev. B **7**, 5212 (1973).

[22] S. Froyen, Phys. Rev. B **39**, 3168 (1989).

[23] J. Moreno and J. M. Soler, Phys. Rev. B **45**, 13891 (1992).

[24] H. J.Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[25] J. C. Phillips, Phys. Rev. **112**, 685 (1958).

[26] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959).

[27] M. L. Cohen and V. Heine, Solid State Physics **24**, 37 (1970).

[28] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[29] G. B. Bachelet, D. R. Hamann, and C. Chiang, Phys. Rev. B **26**, 4199 (1982).

[30] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[31] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[32] M. C. Payne, M. P. Teter, D. C. Allen, T. A. Arias, and J. D. Joannopolous, Rev. Mod. Phys. **64**, 1045 (1992).

[33] E. Polak, *Computational Methods in Optimisation* (Academic Press, New York, 1971).

[34] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, London, 1981).

[35] M. Teter, M. C. Payne, and D. C. Allen, Physical Review B **40**, 12255 (1989).

[36] M. D. Segall, P. J. D. Lindon, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys: Condens. Matter **14**, 2717 (2002).

[37] J. Perdew and K. Burke, Int. J. Quantum Chem. **57**, 309 (1996).

[38] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[39] O. Gunnarsson and R. O. Jones, Solid State Commun. **37**, 249 (1981).

[40] J. G. Harrison, R. A. Heaton, and C. C. Lin, J. Chem. Phys. **78**, 4562 (1983).

[41] R. A. Heaton, J. G. Harrison, and C. C. Lin, Phys. Rev. B **28**, 5992 (1983).

[42] R. A. Heaton and C. C. Lin, J. Phys. C. **17**, 1853 (1984).

[43] A. Svane and O. Gunnarsson, Phys. Rev. Lett **65**, 1148 (1990).

[44] Z. Szotek, W. W. Temmerman, and H. Winter, Phys. Rev. B **47**, 4029 (1993).

[45] J. Hubbard, Proc. R. Soc. London Ser. A **243**, 336 (1957).

[46] D. Pines and P. Nozières, (Benjamin, Reading, Mass., 1966).

[47] M. Gell-Mann and K. A. Brueckner, Phys. Rev. **106**, 364 (1957).

[48] W. J. Carr and A. A. Maradudin, Phys. Rev. **133**, A371 (1964).

[49] E. Wigner, Phys. Rev. **46**, 1002 (1934).

[50] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[51] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[52] J. P. Perdew, Phys. Rev. Lett. **55**, 1665 (1985).

[53] J. P. Perdew, *Electronic Structure of Solids '91* (Akademie Verlag, Berlin, 1991).

[54] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).

[55] A. D. Becke, J. Chem. Phys. **38**, 3098 (1988).

[56] M. Levy and J. P. Perdew, Phys. Rev. A **32**, 2010 (1985).

[57] E. H. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).

[58] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[59] M. Rasolt and D. J. W. Geldart, Phys. Rev. B **34**, 1325 (1986).

[60] J. P. Perdew, Physica B **172**, 1 (1991).

[61] D. C. Langreth and M. J. Mehl, Phys. Rev. B **28**, 1809 (1983).

[62] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Lett. A **146**, 256 (1990).

[63] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[64] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. (E) **78**, 1396 (1997).

[65] S. K. Ghosh and R. G. Parr, Phys. Rev. A **34**, 785 (1986).

[66] A. D. Becke and M. R. Roussel, Phys. Rev. A **39**, 3761 (1989).

[67] E. Proynov, A. Vela, and D. R. Salahub, Chem. Phys. Lett. **230**, 419 (1995).

[68] E. P. E. Ruiz, A. Vela, and D. R. Salahub, Int. J. Quan. Chem. Symp. **29**, 61 (1995).

[69] C. Adamo, M. Ernzerhof, and G. E. Scuseria, J. Chem. Phys. **112**, 2643 (2000).

[70] T. V. Voorhis and G. E. Scuseria, J. Chem. Phys. **109**, 400 (1998).

[71] M. Filatov and W. Thiel, Phys. Rev. A **57**, 189 (1998).

[72] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **111**, 911 (1999).

[73] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, Phys. Rev. Lett. **82**, 2544 (1999).

[74] A. D. Becke, J. Chem. Phys. **112**, 4020 (2000).

[75] E. Proynov, H. Chermette, , and D. R. Salahub, J. Chem. Phys. **113**, 10013 (2000).

[76] D. Boese and N. C. Handy, J. Chem. Phys. **116**, 9559 (2002).

[77] S. Kurth, J. P. Perdew, and P. Blaha, Int. J. Quantum Chem. **75**, 889 (1999).

[78] J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).

[79] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).

[80] A. D. Becke, J. Chem. Phys. **98**, 1372 (1992).

[81] A. D. Becke, J. Chem. Phys. **107**, 8554 (1997).

[82] P. J. Stevens, J. F. Devlin, C. F. Chabolowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[83] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[84] O. Gunnarsson, M. Jonson, and B. I. Lundqvist, Phys. Lett. **59A**, 177 (1976).

[85] O. Gunnarsson, M. Jonson, and B. I. Lundqvist, Solid State Commun. **24**, 765 (1977).

[86] J. A. Alonso and L. A. Girifalco, Phys. Rev. B **17**, 3735 (1978).

[87] O. Gunnarsson, M. Jonson, and B. I. Lundqvist, Phys. Rev. B **20**, 3136 (1979).

[88] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[89] R. C. Grimm and R. G. Storer, J. Comput. Comm. **7**, 134 (1971).

[90] J. B. Anderson, J. Chem. Phys **63**, 1499 (1975).

[91] J. B. Anderson, J. Chem. Phys **65**, 4121 (1976).

[92] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).

[93] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).

[94] A. D. Becke, J. Chem. Phys. **84**, 4524 (1986).

[95] A. D. Becke, J. Chem. Phys. **97**, 9173 (1992).

[96] A. D. Becke, J. Chem. Phys. **96**, 2155 (1992).

[97] M. Körling and J. Häglund, Phys. Rev. B **45**, 13293 (1992).

[98] F. A. Hamprecht, A. J. Cohen, D. J. Tozer, and N. C. Handy, J. Chem. Phys. **109**, 6264 (1988).

[99] J. A. Altmann and N. C. Handy, Phys. Chem. Chem. Phys. **1**, 5529 (1999).

[100] N. J. Wright, R. B. Gerber, and D. J. Tozer, Chem. Phys. Lett. **324**, 206 (2000).

[101] D. A. Preston and B. A. Gilbert, Chem. Phys. Lett. **326**, 537 (2000).

[102] L. A. Curtiss, K. Ragavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).

[103] Q. Zhao, R. C. Morrison, and R. G. Parr, Phys. Rev. A **50**, 2138 (1994).

[104] H. Stoll, C. M. E. Pavlidou, and H. Preuss, Theor. Chem. Acc. **49**, 143 (1979).

[105] J. Hellmann, *Einführung in die Quantenchemie* (Deuticke, Liepzig, 1937).

[106] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[107] P. Pulay, Mol. Phys. **17**, 197 (1969).

[108] S. Kurth, J. P. Perdew, and P. Blaha, Int. J. Quantum Chem. **75**, 889 (1999).

[109] O. Madelung, M. Schulz, and H. Weiss, editors, *Physics of Group IV Elements and III-V Compounds*volume 17 (Springer, Berlin, 1982).

[110] F. D. Murnaghan, Proc. Natl. Acad. Sci. USA **30**, 244 (1944).

[111] J. H. Edgar, editor, (IEE, London, 1994, 1994).

[112] R. A. Felice, J. Trivisonno, and D. E. Schuele, Phys. Rev. B **16**, 5173 (1977).

[113] J. F. Nye, (Clarendon, Oxford, 1985).

[114] A. D. Corso, A. Pasguarello, A. Baldereschi, and R. Car, Phys. Rev. B **53**, 1180 (1996).

[115] M. Fuchs, M. Bockstedte, E. Pehlke, and M. Scheffler, Phys. Rev. B (1998).

[116] M. E. Sherwin and T. J. Drummond, J. Appl. Phys. **69**, 8423 (1991).

[117] C. M. Bertoni, L. Reining, and F. Finocchi, Physica B **185**, 404 (1993).

[118] Y. Juan and E. Kaxiras, Phys. Rev. B **51**, 9521 (1995).

[119] G. Ortiz, Phys. Rev. B **45**, 11328 (1992).

[120] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).

[121] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

[122] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[123] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[124] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[125] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[126] M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10031 (1999).

[127] S. Strite, J. Ruan, Z. Li, A. Salvador, H. Chen, D. J. Smith, W. J. Choyke, , and H. Mork, J. Vac. Sci. Technol. B **9**, 1924 (1991).

[128] D. J. Tozer and N. C. Handy, J. Chem. Phys. **109**, 10180 (1998).

[129] P. J. Wilson, G. Menconi, and D. J. Tozer, (Private communication) .

[130] A. J. Cohen and N. C. Handy, Mol. Phys. **99**, 607 (2001).

[131] G. Menconi and D. J. Tozer, J. Chem. Phys. **114**, 3958 (2001).

[132] D. J. Tozer, (Private communication) .

[133] G. Menconi and D. J. Tozer, (Unpublished).

[134] J. Harris, Phys. Rev. A **29**, 1648 (1984).

[135] G. P. Kerker, Phys. Rev. B **24**, 3468 (1981).

[136] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **30**, 5777 (1984).

[137] F. Manghi, G. Rieglera, C. M. Bertoni, C. Calandra, and G. B. Bachelet, Phys. Rev. B **28**, 6157 (1983).

[138] O. Gunnarsson and R. O. Jones, Phys. Scr. **21**, 394 (1980).

[139] P. García-González, J. E. Alvarellos, E. Chacón, and P. Tarazona, Phys. Rev. B **62**, 16063 (2000).

[140] N. Marzari and D. J. Singh, J. Phys. Chem. Solids **60**, 321 (2000).

[141] D. J. Singh, Phys. Rev. B **48**, 14099 (1993).

[142] L. Fritsche, Phys. Rev. B **33**, 3967 (1986).

[143] J. P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).

[144] P. Krüger and G. Wolfgarten, Solid State Commun. **53**, 885 (1985).

[145] R. Q. Hood, M. Y. Chou, A. J. Williamson, G. Rajagopal, R. J. Needs, and W. M. C. Foulkes, Phys. Rev. Lett. **78**, 3350 (1997).

[146] M. Sadd and M. P. Teter, Phys. Rev. B **64**, 13643 (1996).

[147] M. Sadd and M. P. Teter, J. Mol. Structure (Theochem) **501**, 147 (2000).

[148] O. V. Gritsenko, N. A. Cordero, A. Rubio, L. C. Balbás, and J. A. Alonso, Phys. Rev. A **48**, 4197 (1993).

[149] A. C. Pedroza, Phys. Rev. A **33**, 804 (1986).

[150] J. P. A. Charlesworth, Phys. Rev. B **53**, 12666 (1996).

[151] J. C. Kimball, Phys. Rev. B **7**, 1648 (1973).

[152] A. K. Rajagopal, J. C. Kimball, and M. Banerjee, Phys. Rev. B **18**, 2339 (1978).

[153] J. E. Ortega and F. J. Humpsel, Phys. Rev. B **47**, 2130 (1993).

[154] M. Städele, M. Moukara, J. A. Mejewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10031 (1999).

[155] R. Hulthen and N. G. Nilsson, Solid State Commun. **18**, 1341 (1976).

[156] E. Shirley, X. Zhu, and S. G. Louie, Phys. Rev. B **56**, 6648 (1997).

[157] J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16533 (1996).

[158] M. Ernzerhof and J. P. Perdew, J. Chem. Phys. **109**, 3313 (1998).

[159] R. Q. Hood, M. Y. Chou, A. J. Williamson, G. Rajagopal, and R. J. Needs, Phys. Rev. B **58**, 8972 (1998).

[160] K. S. Singwi, M. P. Tosi, R. H. Land, and A. Sjölander, Phys. Rev. **176**, 589 (1968).

[161] M. Nekovee, W. M. C. Foulkes, and R. J. Needs, Phys. Rev. Lett. **87**, 036401 (2001).

[162] M. Nekovee, W. M. C. Foulkes, A. J. Williamson, G. Rajagopal, and R. J. Needs, Adv. Quantum Chem. **33**, 189 (1999).

[163] M. Nekovee, (Private communication, 2002) .

[164] K. Burke, F. G. Cruz, and K.-C. Lam, J. Chem. Phys. **109**, 8161 (1998).

[165] Y.-H. Kim, I.-H. Lee, S. Nagaraja, J.-P. Leburton, R. Q. Hood, and R. M. Martin, Phys. Rev. B **61**, 5202 (2000).

[166] L. Pollack and J. P. Perdew, J. Phys.: Condens. Matter **12**, 1239 (2000).

[167] P. García-González, Phys. Rev. B **62**, 2321 (2000).

[168] C. Weisbuch and B. Vinter, *Quantum Semiconductor Structures* (Boston, MA: Academic, 1991).

[169] B. Vinter, Phys. Rev. B **20**, 2395 (1979).

[170] T. Ando, B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[171] C. Filippi, C. J. Umrigar, and M. Taut, J. Chem. Phys. **100**, 1290 (1994).

[172] R. Neumann, R. H. Nobes, and N. C. Handy, Mol. Phys. **87**, 1 (1996).

[173] J. P. Perdew and K. Schmidt, Private Communication (2000).