SAND2006-7540

SANDIA REPORT

SAND2006-xxxx
Unlimited Release
Printed November 2006

Beyond the Local Density Approximation: Improving Density Functional Theory for High Energy Density Physics Applications

Normand A. Modine, Alan F. Wright,
Richard P. Muller, Mark P. Sears, Ann E. Mattsson,
and Michael P. Desjarlais

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94-AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2006-xxxx
Unlimited Release
Printed November 2006

Beyond the Local Density Approximation: Improving Density Functional Theory for High Energy Density Physics Applications

Normand A. Modine and Alan F. Wright
Physical, Chemical, & Nano Sciences Center

Richard P. Muller, Mark P. Sears, and Ann E. Mattsson
Computation, Computers, Information and Mathematics Center

Michael P. Desjarlais
Pulsed Power Sciences Center

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185

Abstract

A finite temperature version of ”exact-exchange” density functional theory (EXX) has been implemented in Sandia’s Socorro code. The method uses the optimized effective potential (OEP) formalism and an efficient gradient-based iterative minimization of the energy. The derivation of the gradient is based on the density matrix, simplifying the extension to finite temperatures. A stand-alone all-electron exact-exchange capability has been developed for testing exact exchange and compatible correlation functionals on small systems. Calculations of eigenvalues for the helium atom, beryllium atom, and the hydrogen molecule are reported, showing excellent agreement with highly converged quantum Monte Carlo calculations. Several approaches to the generation of pseudopotentials for use in EXX calculations have been examined and are discussed. The difficult problem of finding a correlation functional compatible with EXX has been studied and some initial findings are reported.

Acknowledgments

We are indebted to Ross Lippert of MIT for his substantial contributions to this work, particularly with regard to the mathematical developments discussed in Chapter 2 and published in Ref. [16]. This work would not have been possible without the combined support of the High Energy Density Physics (HEDP) and Materials Science and Technology (MS&T) LDRD Investment Areas. We would also like to thank Charles Barbour, John Aidun, and Tom Mehlhorn for their early and continued support of this project.

Contents

Appendix

List of Figures

List of Tables

Nomenclature

HEDS High Energy Density Science

DFT Density Functional Theory

OEP Optimized Effective Potential

EXX Exact Exchange

HF Hartree-Fock

LDA Local Density Approximation

PZ A parametrisation of the Ceperley-Alder [35] LDA correlation created by Perdew and Zunger [32]

PW A parametrisation of the Ceperley-Alder [35] LDA correlation created by Perdew and Wang [31]

GGA Generalized Gradient Approximation

PBE A GGA created by Perdew, Burke, and Ernzerhof [33]

PW91 A GGA created by Perdew et al. [34]

RPA+ An enhanced Random Phase Approximation [29]

$\bar{h}$ Planck's constant $= 1.05459 \times 10^{-34}\ J \cdot s$

m_e Electron mass $= 9.1095 \times 10^{-31}\ kg$

e Electron charge $= 1.60219 \times 10^{-19}\ As$

ε_0 Vacuum permittivity $= 8.85419 \times 10^{-12}\ As/Vm$

a_0 Bohr radius $= \bar{h}^2/(m_e(e^2/(4\pi\varepsilon_0))) = 0.529177 \times 10^{-10}\ m$

n Density

Chapter 1

Introduction and Motivation

An integral component of the High Energy Density Science (HEDS) work done at Sandia is the use of advanced modeling codes such as ALEGRA to simulate the complex evolution of materials through solid, liquid, vapor, and plasma phases in HEDS experiments. These codes require accurate equation-of-state (EOS), conductivity, and opacity models if high-fidelity results are to be obtained. A particularly difficult region to characterize is the "warm dense matter" regime that extends from near solid conditions into the vapor dome and temperatures up to several eV. This region of phase space includes the molecular-to-atomic dissociation phase for dense hydrogen and its isotopes, and the metal-insulator transition for liquid metals. In recent years, much progress has been made in the understanding and accurate modeling of these regimes through the use of quantum molecular dynamics based on finite temperature density-functional theory (FT-DFT), which enables the calculation of manifestly consistent EOS, conductivities, and low-energy opacities [1, 2, 3, 4]. Although the application of density functional methods to high energy density science is rather new, DFT is a powerful and commonly used tool in Material Science and Technology (MS&T) research. Applications include surface science, the properties of water, the study of defects in materials, insulators, complex materials, optical ceramics, and oxides.

The commonly used DFT codes of today employ relatively simple but efficient, explicit functionals of the density, categorized as local density or generalized gradient approximations (LDA and GGA), respectively. These functionals have proven to give quite accurate results for many properties of interest. However, as numerical methods have evolved, and large systems and complex problems have been studied with high precision, the deficiencies of these functionals have become the limiting factor to needed improvements in accuracy. It is well-known that LDA and GGA functionals typically underestimate the energy gap between occupied (valence) and unoccupied (conduction) bands by 1 to 2 eV, shifting behavior more towards a metal and underestimating metallization densities and pressures. In HEDS applications, when the temperature is comparable to or smaller than the gap, the conductivity and the low-energy opacities will be significantly increased. In these HEDS applications it is often the case that the temperature is high enough to thermally populate many of the conduction bands, and therefore corrections to the band gap have a corresponding effect on the pressure and energy, as well as the conductivity and optical properties. The only way to self-consistently improve calculations where the erroneous gap influences the results is to develop and use an improved functional. Improvements to the accuracy of the exchange functional will provide a significantly enhanced tool for MS&T research as well. For example, since defect formation energies and charges depend on the position of defect levels in the gap, improve-

ment in band gaps will help in identifying defects and predicting defect populations, kinetics, and properties during growth and processing, and following radiation damage in materials of interest to Sandia (e.g., alumina).

Accordingly, we have developed and implemented a much more advanced treatment of the exchange-correlation functional into Sandia's state-of-the art plane-wave DFT computational framework, Socorro. This approach centers on using the known exact expression for exchange in the functional. The exact expression is a highly non-local, implicit functional of the density (an explicit functional of the Kohn-Sham orbitals). Doing this leaves the evaluation of the smaller correlation contribution as the only remaining approximation. Several groups have already performed small-scale investigations of this sort. Use of exact exchange (EXX) is found to essentially solve the band gap and metallization problems. It remains an important and challenging problem to combine EXX-DFT with molecular dynamics for systems of several tens to hundreds of atoms, but the potential return is broadly important and should have enduring impact in computational material science.

Chapter 2

Mathematics of the optimized effective potential (OEP)

Introduction

The Kohn-Sham Density Functional Theory (DFT) [5] has become one of the most powerful tools for understanding and predicting the properties of materials. DFT has been applied to an ever increasing number of different types of systems and phenomena, and the results have frequently been remarkably useful. Nevertheless, the accuracy of the results remains an important issue for many potential applications of DFT. The main source of error in DFT calculations is the use of an approximate expression for the exchange-correlation energy, E_{XC}. Such an approximation is necessary for practical calculations, but improving the quality of the approximation, and hence, the accuracy of the calculations, is of great interest. Conventional variants of DFT, such as LDA and GGA, take E_{XC} to be an explicit functional of the electronic density. Since the noninteracting Kohn-Sham orbitals are implicit functionals of the electronic density [6], expressions for E_{XC} that explicitly depend on the Kohn-Sham orbitals are also consistent with the DFT framework. An important example of such a functional is the functional used in the exact-exchange approximation[7, 8, 9, 10, 11, 12, 13]. An explicit dependence on the orbitals allows approximate E_{XC} expressions to capture physical behaviors of the exact Kohn-Sham E_{XC} that can not be practically incorporated in an expression that is an explicit function of only the electronic density. One example is the absence of self-interaction in the exact Kohn-Sham energy. Another example is the complex, non-local behavior of the exact exchange energy.

The difficulty in using an E_{XC} expression that is explicitly dependent on the orbitals is that it is impossible to straightforwardly take the functional derivative of E_{XC} with respect to the electronic density. Therefore, standard self-consistent methods of minimizing the energy with respect to the density can not be used. The solution to this problem is provided by the Optimized Effective Potential (OEP) formalism. Since the energy is a functional of the Kohn-Sham orbitals, and the orbitals are solutions of the Kohn-Sham equation for some local potential, the energy can be viewed as a functional of the potential. The OEP is defined to be the potential that minimizes the energy. This minimization with respect to the potential is equivalent to the usual minimization with respect to the density. Traditionally, the OEP has been calculated by solving the OEP integral equation, in which the gradient of the energy with respect to the potential is set to zero [7, 8, 9, 11], or by

directly evaluating and inverting a response function [10, 12, 13].

Two recent papers have proposed calculation of the OEP by means of an iterative minimization of the energy [14, 15]. Hyman, Stiles, and Zangwill used Lagrange multiplier methods to derive an expression for the gradient of the energy with respect to the potential and proposed using this gradient to minimize the energy iteratively. Kummel and Perdew derived a nearly identical expression and, although they did not claim that this expression gives the gradient, they noted that it provides a good update to the potential during an iterative minimization. In this paper, we present a new derivation of the gradient based on the density matrix. Our work goes beyond the previous papers in the following ways: (1) We believe that our derivation is particularly transparent, and therefore, it demonstrates that this expression is, in fact, the correct gradient. (2) The previous work assumed a negligible electronic temperature. Since our derivation is based on the density matrix, it is easily extended to finite temperatures, where the orbitals are partially occupied.

One of the most exciting recent applications of DFT has been high energy density physics. In this application, electronic temperatures that are substantial compared to the band gaps of typical semiconductors are common. This makes the results sensitive to the band gap, which is too small in the standard versions of DFT. Therefore, the capability of performing calculations with advanced functionals that have explicit dependence on the orbitals at non-zero temperature is particularly exciting for high energy density physics applications.

As an alternative to iterative minimization of the energy using the gradient, it is possible, in principle, to find the OEP by solving the equation in which the gradient is set to zero. Therefore, our work provides the finite temperature equivalent of the standard OEP equation, giving the correct necessary condition for local optimality.

In this article, we derive the OEP method in a finite temperature regime by considering the perturbation of the density matrix resulting from a perturbed Hamiltonian. The gradient will reduce to a combination of orbital shifts as one sees in the zero temperature limit plus some corrections which come from the finite temperature. In section 2, we begin with a mathematical discussion of the perturbation theory of analytic functions of Hermitian operators. After a short review of density functional theory, we apply the results of Section 2 to the density matrix ρ viewed as a function of the Kohn-Sham Hamiltonian H, and thereby derive a finite temperature OEP equation in terms of H and ρ. This motivates the subsequent section, which describes the gradient expression in *orbital form*. We conclude with some computational results demonstrating the accuracy of the method. A more streamlined version of these results with alternative derivations of some of the expressions has been published recently [16].

The perturbation theory of matrix-analytic functions

Let $f(x)$ be an analytic function of x and $f(A)$ be the extension of f to a matrix-analytic function (see [17], chapter 6) on some algebra of Hermitian operators with a finite (or countable) spectrum.

Thus,

$$[A, f(A)] = Af(A) - f(A)A = 0 \tag{2.1}$$
$$f_i = f(a_i) \tag{2.2}$$

where the a_i are the eigenvalues of A and the f_i are the eigenvalues of $f(A)$.

For an unconstrained variation $A \to A + \delta A$ the variations of (2.1) and (2.2) are

$$[\delta A, f(A)] + [A, \delta f(A)] = 0 \tag{2.3}$$
$$\delta f_i = f'(a_i)\delta a_i. \tag{2.4}$$

In a basis where A is diagonal ($a_i = A_{ii}$, $f(a_i) = [f(A)]_{ii}$), (2.3) and (2.4) become,

$$(a_i - a_j)[\delta f(A)]_{ij} = (f(a_i) - f(a_j))\delta A_{ij} \tag{2.5}$$
$$\delta [f(A)]_{ii} = f'(a_i)\delta A_{ii}. \tag{2.6}$$

Thus (2.3) and (2.4) appear to be sufficient to determine $\delta f(A)$ in terms of $A, \delta A$.

We may, somewhat informally, write the result as one equation

$$[\delta f(A)]_{ij} = \frac{f(a_i) - f(a_j)}{a_i - a_j}\delta A_{ij} \tag{2.7}$$

where it is understood that we treat $\frac{f(a_i)-f(a_j)}{a_i-a_j}$ as a *divided difference*, taking the limit as $a_i \to a_j$.

A more rigorous proof of these results can be made in the following theorem, which also makes clear what happens in the presence of a repeated eigenvalue.

Theorem 2.0.1 *The expansion of $f(A+\delta A) - f(A) = \delta f(A)$ to first order in δA is given by*

$$[\delta f(A)]_{ij} = \lim_{\varepsilon \to 0} \frac{f(a_i + \varepsilon) - f(a_j)}{a_i - a_j + \varepsilon}\delta A_{ij}$$

where the matrix elements are taken in an basis in which A is diagonal.

Proof: Since f is analytic, it suffices to prove this theorem for $f(x) = x^k$ and extend by linearity.

$$\delta f(x) = \sum_{m+n=k-1} A^m \delta A A^n$$

in a diagonal basis,

$$\begin{aligned}
[\delta f(x)]_{ij} &= \sum_{m+n=k-1} a_i^m a_j^n \delta A_{ij} \\
&= \sum_{m+n=k-1} \lim_{\varepsilon\to 0} (a_i+\varepsilon)^m a_j^n \delta A_{ij} \\
&= \lim_{\varepsilon\to 0} \frac{(a_i+\varepsilon)^k - a_j^k}{(a_i+\varepsilon)-a_j} \delta A_{ij}
\end{aligned}$$

and the summation is interchanged with the limit. □

Note that one could clearly obtain higher order derivatives in terms of higher order divided differences via the same approach. To simplify forthcoming derivations, we omit ε's and limits, with the understanding that appropriate limits are to be taken for divided differences of the form $\frac{f(x)-f(y)}{x-y}$. We may interpret (2.7) as the equation which specifies the action of the *Jacobian*, $\frac{\partial f(A)}{\partial A}$ on an arbitrary Hermitian operator of the mapping f.

Practical application of the Jacobian

An A-diagonalizing basis might not be a convenient means to compute the application of the Jacobian to an arbitrary variation. We present here a short digression on how such computations can be carried out iteratively, without diagonalizing A.

We will be interested in applications of linear operators to Hermitian (or anti-Hermitian) matrices. To avoid some of the confusion entailed in *operators of operators* discussions, we introduce some notation, which we hope is clarifying. We denote the application of a linear operator on a matrix with brackets, $L[A]$. In terms of indices we may write this as $L[A]_{ij} = \sum_{kl} L_{ijkl} A_{kl}$. For Hermitian matrices, we write $\langle X,Y\rangle = \mathrm{tr}\{XY\}$, and it is well known that this is a non-degenerate inner product on the vector space of Hermitian matrices. For a vector subspace of Hermitian matrices, A, we let $A^\perp = \{B : \forall X \in A, \langle X,B\rangle = 0\}$.

We denote the linear action of a commutator $\mathrm{ad}_X[Y] = -\mathrm{ad}_Y[X] = [X,Y]$.

Lemma 2.0.2

$$\langle X, ad_Y[Z]\rangle = -\langle ad_Y[X], Z\rangle .$$

Proof: Applying the trace identity $\mathrm{tr}\{AB\} = \mathrm{tr}\{BA\}$,

$$\begin{aligned}
\mathrm{tr}\{X(YZ-ZY)\} &= \mathrm{tr}\{XYZ - YXZ\} \\
&= -\mathrm{tr}\{(YX-XY)Z\}
\end{aligned}$$

□

The *centralizer* of X is the set $C_X = \{Y : [X,Y] = 0\}$. C_X is the nullspace of ad_X.

Lemma 2.0.3 *$C_X^\perp$ is the range of ad_X.*

Proof: For any Y and some $Z \in C_X$, by lemma 2.0.2, $\langle \mathrm{ad}_X(Y), Z\rangle = -\langle Y, \mathrm{ad}_X(Z)\rangle = 0$, thus $\mathrm{Range}\{\mathrm{ad}_X\} \subset C_X^\perp$.

From elementary dimension counting,

$$\begin{aligned} \dim\{\mathrm{Range}\{\mathrm{ad}_X\}\} + \dim\{\mathrm{Null}\{\mathrm{ad}_X\}\} &= \dim\{C_X\} + \dim\{C_X^\perp\} \\ \dim\{\mathrm{Range}\{\mathrm{ad}_X\}\} + \dim\{C_X\} &= \dim\{C_X\} + \dim\{C_X^\perp\} \\ \dim\{\mathrm{Range}\{\mathrm{ad}_X\}\} &= \dim\{C_X^\perp\} \end{aligned}$$

where $\dim\{C_X\} = \dim\{C_X\}$ obtains the last line. □

Let $X_{ij} = \frac{f(a_i)-f(a_j)}{a_i-a_j} Y_{ij}$ in an A-diagonalizing basis. Then according to (2.1) and (2.3),

$$\mathrm{ad}_A[X] = \mathrm{ad}_{f(A)}[Y]. \tag{2.8}$$

The operator, ad_A, has a non-trivial nullspace. However, $C_A \subset C_{f(A)}$ implies $\mathrm{Range}\{\mathrm{ad}_{f(A)}\} \subset \mathrm{Range}\{\mathrm{ad}_A\}$, by lemma 2.0.3, thus equation (2.8) has a unique solution.

Since we are dealing with a linear space (albeit of matrices), with a linear operator and an inner product, we can use a Krylov-based iterative solver to solve (2.8) (e.g. conjugate gradient or MINRES [18, 19]) with some initial guess, X_0, yielding

$$X = X_0 + c_1 \mathrm{ad}_A[X_0] + c_2 \mathrm{ad}_A^2[X_0] + \cdots$$

with $X - X_0 \in C_A^\perp$.

By (2.4), we additionally require $X_{ii} = f'(a_i) Y_{ii}$. For example, if we take

$$X_0 = \frac{1}{2}\left(f'(A)Y + Y f'(A)\right) + X_1,$$

where $X_1 \in C_A^\perp$ is arbitrary, then the iterative solution of (2.8) will be correct, i.e. $X = \delta f(A)$. In the remainder of this article we take $X_1 = 0$.

Density functional theory review

Let ρ be a density matrix (Hermitian), K be the kinetic energy operator, V_I be the ionic (and external) potential, with $E_{HXC}(\rho)$ the Hartree, exchange, and correlation energy. With $S(P) = -\mathrm{tr}\{\rho\log(\rho)+(I-\rho)\log(I-\rho)\}$ as the entropy expression, the variational energy is

$$E(\rho) = \mathrm{tr}\{\rho(K+V_I)\} + E_{HXC}(\rho) - \frac{1}{\beta}S(\rho). \tag{2.9}$$

The unconstrained derivative is

$$\frac{\partial E}{\partial \rho} = K + V_I + \frac{\partial E_{HXC}}{\partial \rho} + \frac{1}{\beta}\log(\rho(I-\rho)^{-1}). \tag{2.10}$$

The Kohn-Sham Hamiltonian is given by $H = K + V_I + V$ where V is the self-consistent potential (to be determined). In the Kohn-Sham DFT, ρ is the minimizer of $\mathrm{tr}\{\rho H\} - \frac{1}{\beta}S(\rho)$ with $\mathrm{tr}\{\rho\} = n$, which is equivalent to the conditions,

$$\rho = \frac{1}{1+e^{\beta(H-\mu I)}} = f_\beta(H-\mu I) \tag{2.11}$$

$$\mathrm{tr}\{\rho\} = n \tag{2.12}$$

for some chemical potential μ. Thus, we can consider ρ to be parametrized by two unknowns V and μ with two relations (2.11) and (2.12). Note: one could absorb μ into V, but we find it advantageous to keep it distinct in its role as a Lagrange multiplier.

With ρ satisfying these relations, the energy differential simplifies

$$\begin{aligned}\frac{\partial E}{\partial \rho} &= K + V_I + \frac{\partial E_{HXC}}{\partial \rho} + \frac{1}{\beta}\log(\rho(I-\rho)^{-1}) \\ &= K + V_I + \frac{\partial E_{HXC}}{\partial \rho} - (H-\mu I) \\ &= \frac{\partial E_{HXC}}{\partial \rho} - (V-\mu I).\end{aligned}$$

Finite temperature OEP with density operators

From section 2, the density matrix is related to the Kohn-Sham Hamiltonian, H, by (2.11) and (2.12). Let ε_i be the eigenvalues of H, and let $\omega_i = f_\beta(\varepsilon_i - \mu)$ be the eigenvalues of ρ. The divided

differences can be stably computed with the formula

$$\frac{f_\beta(x)-f_\beta(y)}{x-y} = -\frac{e^{\beta(x+y)/2}}{(1+e^{\beta x})(1+e^{\beta y})}\left(\frac{\sinh(\beta(x-y)/2)}{(x-y)/2}\right)$$

where a test for $x=y$ is required for the evaluation of the last factor. We note in particular that $\frac{f_\beta(x)-f_\beta(x)}{x-x} = \frac{d}{dx}f_\beta(x) = -\beta f_\beta(x)(1-f_\beta(x))$.

Let $E(\rho)$ be a function of a density matrix, ρ. We can implicitly define $E(H) = E(\rho(H,\mu(H)))$. Formally varying $E(H)$,

$$\delta E(H) = \mathrm{tr}\left\{\frac{\partial E(\rho(H,\mu))}{\partial\rho}\delta\rho\right\}. \tag{2.13}$$

in an H-diagonalizing basis,

$$\delta\rho_{ij} = \frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}\left(\delta H_{ij}-\delta\mu\,\delta_{ij}\right). \tag{2.14}$$

By (2.12), the trace of $\delta\rho$ vanishes,

$$\begin{aligned}\delta\mu &= \frac{\sum_{ij}\delta_{ij}\frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}\delta H_{ij}}{\sum_{ij}\delta_{ij}\frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}}\\ &= \frac{\mathrm{tr}\{\rho(I-\rho)\delta H\}}{\mathrm{tr}\{\rho(I-\rho)\}}.\end{aligned}$$

Thus, in an H-diagonalizing basis,

$$\begin{aligned}\delta E &= \sum_{ij}\frac{\partial E}{\partial\rho_{ij}}\frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}\left(\delta H_{ij}-\delta\mu\,\delta_{ij}\right)\\ &= \sum_{ij}\frac{\partial E}{\partial\rho_{ij}}\frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}\left(\delta H_{ij}-\frac{\sum_k\omega_k(1-\omega_k)\delta H_{kk}}{\sum_k\omega_k(1-\omega_k)}\delta_{ij}\right)\\ &= \sum_{ij}\left(\frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}\frac{\partial E}{\partial\rho_{ij}}-\delta_{ij}\frac{\omega_i(1-\omega_i)}{\sum_k\omega_k(1-\omega_k)}\sum_{pq}\delta_{pq}\frac{\omega_p-\omega_q}{\varepsilon_p-\varepsilon_q}\frac{\partial E}{\partial\rho_{pq}}\right)\delta H_{ij}\end{aligned}$$

and the gradient is therefore

$$\frac{\partial E}{\partial H_{ij}} = \frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}\frac{\partial E}{\partial \rho_{ij}} - \delta_{ij}\frac{\omega_i(1-\omega_i)}{\sum_k \omega_k(1-\omega_k)}\sum_{pq}\delta_{pq}\frac{\omega_p-\omega_q}{\varepsilon_p-\varepsilon_q}\frac{\partial E}{\partial \rho_{pq}} \tag{2.15}$$

$$\frac{\partial E}{\partial H} = \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho}\right] - \frac{\rho(I-\rho)}{\mathrm{tr}\{\rho(I-\rho)\}}\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho}\right]\right\} \tag{2.16}$$

$$= \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho} - \frac{\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho}\right]\right\}}{\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}[I]\right\}}I\right] \tag{2.17}$$

where $\frac{\Delta\omega}{\Delta\varepsilon}[\cdot]$ stands for the Jacobian, $\frac{\omega_i-\omega_j}{\varepsilon_i-\varepsilon_j}$, in a general basis. $\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}[I]\right\} = -\beta\mathrm{tr}\{\rho(I-\rho)\}$, though we will keep it as it is in (2.17) to make the tracelessness of $\frac{\partial E}{\partial H}$ more manifest.

The application of the Jacobian, to obtain $\frac{\partial E}{\partial H} = \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho} - \frac{\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho}\right]\right\}}{\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}[I]\right\}}I\right]$, can be done by iteratively solving

$$\left[H, \frac{\partial E}{\partial H}\right] = \left[\rho, \frac{\partial E}{\partial \rho}\right] \tag{2.18}$$

with initial guess

$$\left(\frac{\partial E}{\partial H}\right)_0 = -\frac{1}{2}\beta\left(\rho(I-\rho)\frac{\partial E}{\partial \rho} + \frac{\partial E}{\partial \rho}\rho(I-\rho)\right) + \beta\frac{\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho}\right]\right\}}{\mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}[I]\right\}}\rho(I-\rho). \tag{2.19}$$

Note: $\lim_{\beta\to\infty}\beta\rho(I-\rho) \propto \delta(H-\mu I)$, a delta function on the spectrum of H. Thus in the low temperature limit, only the eigenvalues of H near μ contribute to $\left(\frac{\partial E}{\partial H}\right)_0$.

To obtain an OEP gradient, we restrict the variability of H to $H = H_0 + V$ where $H_0 = K + V_I$ is fixed and V is a local operator. The gradient is then

$$\frac{\partial E}{\partial V(r)} = \sum_{ij}\phi_i(r)\frac{\partial E}{\partial H_{ij}}\phi_j^*(r) \tag{2.20}$$

where $\phi_i(r)$ is the eigenvector of H with eigenvalue ε_i in the position representation.

Finite temperature OEP with orbitals

Instead of representing the density as an operator, it is often more practical to express ρ in terms of an incomplete basis of partially occupied orbitals. Let $\phi_1, \phi_2, \ldots$ be a complete eigenbasis of H sorted non-decreasingly in eigenvalue. Let N be sufficiently large that $\omega_{i>N} \sim 0$. Then we may write truncate the basis so that

$$\rho = \phi \Omega \phi^\dagger \tag{2.21}$$

where $\phi = \begin{bmatrix}\phi_1 & \phi_2 & \cdots & \phi_N\end{bmatrix}$, with Λ the diagonal matrix of eigenvalues and Ω the diagonal matrix with entries $\omega_1, \ldots, \omega_N$ (i.e. $\Omega = f_\beta(\Lambda - \mu I)$). Note, N will usually be much smaller than the number of primitive basis functions, so ϕ will be a rectangular matrix with orthonormal columns, i.e. $\phi^\dagger \phi = I$ (the $N \times N$ identity) and $\phi\phi^\dagger$ is the orthogonal projector onto the span of the ϕ_i (and hence commutes with H).

Let χ and ζ be given by

$$\begin{aligned}
\chi &= \mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho}\right]\right\} = \sum_{i\geq 1} -\beta\omega_i(1-\omega_i)\phi_i^\dagger \frac{\partial E}{\partial \rho}\phi_i \\
\zeta &= \mathrm{tr}\left\{\frac{\Delta\omega}{\Delta\varepsilon}[I]\right\} = \sum_{i\geq 1} -\beta\omega_i(1-\omega_i)
\end{aligned}$$

and let $\bar{E}$ and $\bar{J}$ be $N \times N$ matrices given by

$$\bar{E} = \phi^\dagger \frac{\partial E(\phi\Omega\phi^\dagger)}{\partial \rho}\phi \qquad \left(\text{i.e. } \bar{E}_{ij} = \frac{\partial E(\phi\Omega\phi^\dagger)}{\partial \rho_{ij}}\right) \tag{2.22}$$

$$\bar{J} = \phi^\dagger \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho} - \frac{\chi}{\zeta}I\right]\phi \qquad \left(\text{i.e. } \bar{J}_{ij} = \frac{\omega_i - \omega_j}{\varepsilon_i - \varepsilon_j}\left(\bar{E}_{ij} - \frac{\chi}{\zeta}\delta_{ij}\right)\right). \tag{2.23}$$

Note that by these definitions, $[\Omega, \bar{E}] = [\Lambda, \bar{J}]$.

Since $\omega_{i>N} \sim 0$, the expressions for χ and ζ can likewise be truncated,

$$\begin{aligned}
\chi &= \sum_{1\leq i\leq N} -\beta\omega_i(1-\omega_i)\bar{E}_{ii} = -\beta\,\mathrm{tr}\{\Omega(I-\Omega)\bar{E}\} \\
\zeta &= \sum_{1\leq i\leq N} -\beta\omega_i(1-\omega_i) = -\beta\,\mathrm{tr}\{\Omega(I-\Omega)\}
\end{aligned}$$

and by (2.15)

$$(I - \phi\phi^\dagger)\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho} - \frac{\chi}{\zeta}I\right](I - \phi\phi^\dagger) = 0,$$

thus, we may write (2.16) as

$$\frac{\partial E}{\partial H} = \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho} - \frac{\chi}{\zeta} I\right] = \phi\psi^\dagger + \psi\phi^\dagger, \tag{2.24}$$

where $\phi^\dagger\psi = \psi^\dagger\phi = \frac{1}{2}\bar{J}$. This gives an orbital form of equation (2.20),

$$\frac{\partial E}{\partial V(r)} = \sum_{1\le i\le N} \phi_i(r)\psi_i^*(r) + \psi_i(r)\phi_i^*(r) \tag{2.25}$$

which is similar to the equation derived in numerous sources in the OEP literature [14, 15], with a modification of the ψ to accommodate the finite temperature regime.

It remains to solve for ψ, which we may decompose as $\psi = \psi_\perp + \frac{1}{2}\phi\bar{J}$, where $\phi^\dagger\psi_\perp = 0$. We can derive an equation for $\psi_\perp$, by multiplying equation (2.18) on the left by the projector $I - \phi\phi^\dagger$ (which commutes with H) and on the right by ϕ and employing (2.24),

$$(I-\phi\phi^\dagger)\left[H, \frac{\partial E}{\partial H}\right]\phi = (I-\phi\phi^\dagger)\left[\rho, \frac{\partial E}{\partial \rho}\right]\phi \tag{2.26}$$

$$(I-\phi\phi^\dagger)(H\psi - \psi\Lambda) = -(I-\phi\phi^\dagger)\frac{\partial E}{\partial \rho}\rho\phi \tag{2.27}$$

$$H\psi_\perp - \psi_\perp\Lambda = -(I-\phi\phi^\dagger)\frac{\partial E}{\partial \rho}\phi\Omega. \tag{2.28}$$

The LHS and RHS of (2.28) are orthogonal to ϕ, by construction. Thus we have a well-defined equation for $\psi_\perp$. An iterative method can thus be used to solve for $\psi_\perp$ without any special initialization beyond $\phi^\dagger\psi_\perp = 0$.

Computational results

In order to test the above approach, it was implemented in the orbital representation within the Socorro electronic structure software using a plane wave basis set and norm-conserving pseudopotentials. The conjugate gradient algorithm was used to solve the linear systems involved in the evaluation of the gradient. Using this algorithm, the computational cost of solving the set of linear systems determining $\psi_\perp$ is comparable to the cost of solving the Kohn-Sham eigenproblem for ϕ. Therefore, each gradient evaluation is approximately as computationally expensive as one step of the self-consistency loop in a standard DFT code.

For traditional approximations to the exact DFT, such as the Local Density Approximation (LDA) and Generalized Gradient Approximation (GGA), E_{HXC} is an explicit functional of the electronic

density, which is the diagonal of the density matrix ρ in a position representation. In this case, $\frac{\partial E_{HXC}}{\partial \rho}$ has the form of a local potential operator V_{HXC}, and the energy minimum occurs at self-consistency, i.e., when $V = V_{HXC}$. In this case, the OEP is the self-consistent potential, and the results of our iterative minimization approach can be compared directly to well-tested results obtained from conventional self-consistent methods. Therefore, we have tested our OEP approach by applying it to LDA calculations.

Our test system consists of a two atom unit cell of silicon in the diamond structure. We used a 20 Rydberg plane-wave cutoff and a $2 \times 2 \times 2$ Monkhorst-Pack k-point sampling. This k-point sampling does not give a converged total energy, but this is not an issue for the purpose of testing our approach. Two electronic temperatures were used: (1) Room Temperature ($k_B T = 25.67$ meV), and (2) High Temperature ($k_B T = 1.0$ eV).

In order to test the correctness of the our gradient, we used the finite difference approach. During each of a series of steps, the value of V at each point on a real-space grid was varied by a small ($o(10^{-4})$) random perturbation $\Delta V(r)$. During this random walk, the energy and the gradient were evaluated at each step. A linear approximation to the change in energy during each step is given by

$$\Delta E \approx \int \frac{\partial E}{\partial V(r)} \Delta V(r) dr. \tag{2.29}$$

For the small steps taken in this test, we would expect this linear approximation to be accurate if the gradient is accurate. Therefore, we can compare this predicted energy change to the actual energy change observed during the random walk. The results of this comparison for the high temperature case are shown in Fig. 4.1. Since the step direction is random, this represents a very stringent test of the accuracy of the gradient, and we believe that the excellent agreement between the predicted and actual energy changes demonstrates that our approach gives an accurate gradient, even at large electronic temperatures.

The OEP is found by using the gradient to iteratively minimize the energy. We implemented this minimization using Chebyshev acceleration on the fixed point equation $x_{i+1} = x_i + \tau \nabla f(x_i)$ for some fixed τ empirically chosen. The convergence of the energy of our test system during this process is shown in Fig. 2.2. The errors in the energy were evaluated by comparing the energies obtained during the iterative minimization to the result of a highly converged self-consistent calculation. The convergence demonstrates that the iterative OEP and self-consistent approaches give the same result, as would be expected for the LDA energy functional. The convergence is only weakly dependent on the electronic temperature. The asymptotic rate of convergence obtained in the iterative OEP approach is not as rapid as the highly optimized mixing methods typically used in self-consistent calculations, but a reasonable accuracy for practical purposes (10^{-4} Ry.) can be obtained easily.

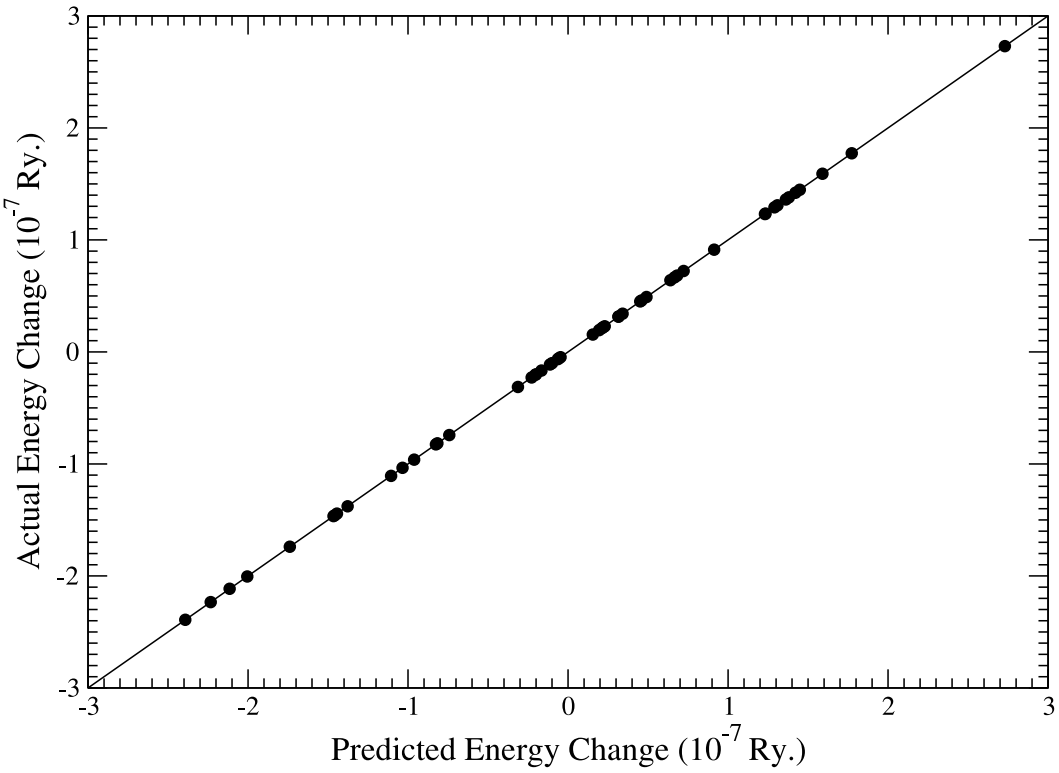


Figure 2.1. A comparison of the energy change predicted from the gradient and the actual energy change observed during a random walk in the potential. The filled circles are the calculated values. The solid line is a guide to the eye representing perfect agreement.

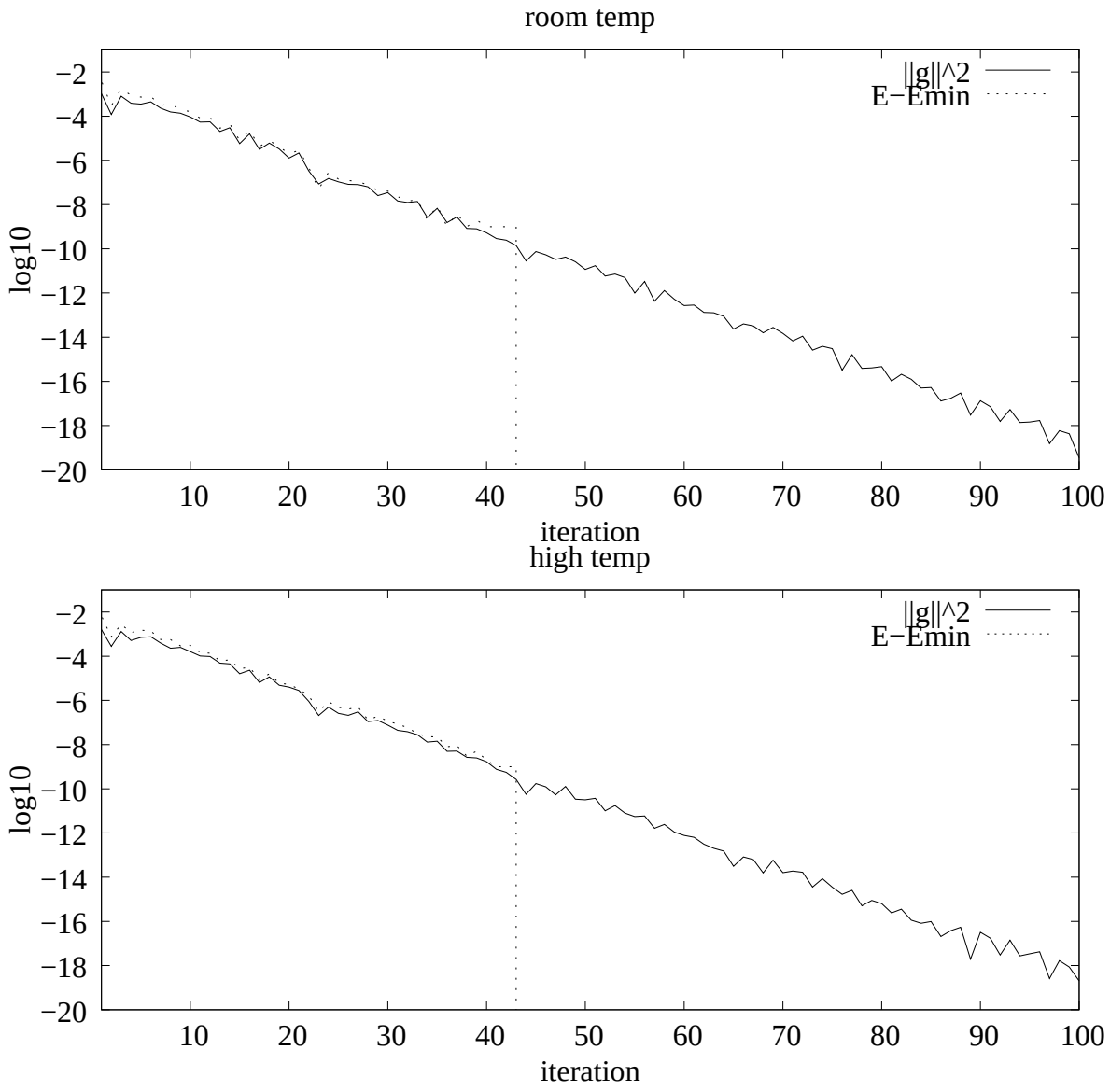


Figure 2.2. The error in the energy, as well as the square-norm of gradient, during the iterative minimization on a log scale.

Finite Basis Set Issues

It has recently been noted [20] that the exact exchange OEP optimization problem seems to suffer from certain pathologies which lead to either non-uniqueness of the effective potential at optimality or the equivalence of the calculated OEP density and energy to that of the corresponding Hartree-Fock quantities.

The argument can be summarized as follows. Let P be any two-point Hermitian function (i.e. $P(x_1,x_2) = \bar{P}(x_2,x_1)$). The total energy functional, in either the Hartree-Fock or exchange-only KS schemes, is

$$E[P] = \int \delta(x-x') \left[-\frac{1}{2}\nabla_x^2 + v_{\text{ext}}(x) + \frac{1}{2}v_J(x) \right] P(x,x') dx dx' + E_X$$

where

$$\begin{aligned} v_J(x) &= \int \frac{P(x',x')}{|x-x'|} dx' \\ E_X &= -\frac{1}{2} \int \frac{|P(x,x')|^2}{|x-x'|} dx dx'. \end{aligned}$$

For the Hartree-Fock method, $E[P]$ is minimized over all P of the form $P = \sum_{i=1}^{n_e} \phi_i(x_1)\bar{\phi}_i(x_2)$ where $\langle \phi_i, \phi_j \rangle = \delta_{ij}$. The local optimality conditions for HF are $\left[P, \hat{H}^{HF}\right] = 0$, where

$$\hat{H}^{HF} = -\frac{1}{2}\nabla^2 + v_{\text{ext}} + v_J + \hat{K}$$

where

$$\hat{K}\phi(x) = \int \frac{P(x',x)}{|x-x'|} \phi(x') dx'.$$

For exact exchange-only Kohn-Sham, $E[P]$ is also optimized with the added condition that the ϕ_i are eigenfunctions of

$$\hat{H}^{KS} = -\frac{1}{2}\nabla^2 + v_{\text{ext}} + v_J + v_X$$

for some local potential v_X, which implies $\left[P, \hat{H}^{KS}\right] = 0$. Since this merely restricts the search space for optimizing $E[P]$, the exact exchange-only KS solution cannot have a lesser energy than the exact HF solution.

The paper by Staroverov et al [20] explores seemingly paradoxical results where, in certain finite bases, it is possible to construct a v_X such that $\left[P,\hat{H}^{KS}\right]=0$ (i.e. P satisfies the exact exchange-only KS constraints) when P is the optimum of the associated Hartree-Fock problem. Thus the minimum exact exchange-only KS energy coincides with that of HF and P minimizes of both problems.

The essence of the argument is the observation that $\hat{H}^{KS}-\hat{H}^{HF}=v_X-\hat{K}$, and hence $\left[P,\hat{H}^{KS}\right]=0$ if $\left[P,\hat{H}^{HF}\right]=0$ and $[P,v_X]=\left[P,\hat{K}\right]$. Taking P to be the HF minimizer gives the first condition, and a the choice of basis functions provides the second. Let $a_\mu(x)$ be real basis functions for the orbitals, thus $P(x_1,x_2)=\sum_{\mu_1,\mu_2}P^{\mu_1\mu_2}a_{\mu_1}(x_1)a_{\mu_2}(x_2)$, and $b_\nu(x)$ be real basis functions for the potentials, $v_X=\sum_\nu v_X^\nu b_\nu(x)$. Then the relevant products are

$$\begin{aligned}
(v_XP)(x_1,x_2) &= \sum_{\mu_1\mu_2\nu} v_X^\nu P^{\mu_1\mu_2} b_\nu(x_1)a_{\mu_1}(x_1)a_{\mu_2}(x_2)\\
(\hat{K}P)(x_1,x_2) &= \sum_{\mu'\mu_2} P^{\mu'\mu_2}\left(\int\frac{P(x',x_1)}{|x_1-x'|}a_{\mu'}(x')dx'\right)a_{\mu_2}(x_2)\\
&= \sum_{\mu_1\mu\mu'\mu_2} P^{\mu'\mu_2}\left(\int\frac{P^{\mu\mu_1}a_\mu(x')a_{\mu_1}(x_1)}{|x_1-x'|}a_{\mu'}(x')dx'\right)a_{\mu_2}(x_2)\\
&= \sum_{\mu_1\mu\mu'\mu_2} P^{\mu\mu_1}P^{\mu'\mu_2}\left(\int\frac{a_\mu(x')a_{\mu'}(x')}{|x_1-x'|}dx'\right)a_{\mu_1}(x_1)a_{\mu_2}(x_2).
\end{aligned}$$

Clearly, we can have $v_XP=\hat{K}P$ by

$$\begin{aligned}
b_{\nu(\mu,\mu')}(x) &= \int\frac{a_\mu(x')a_{\mu'}(x')}{|x-x'|}dx'\\
v_X^{\nu(\mu,\mu')} &= \frac{P^{\mu\mu_1}P^{\mu'\mu_2}}{P^{\mu_1\mu_2}}
\end{aligned}$$

for some appropriate choice of index map, $\nu(\mu,\mu')$ (generically, $P^{\mu_1\mu_2}$ is non-vanishing). Staroverov et al acknowledge that this construction reveals that for a generic basis of wavefunction, $\{a_\mu\}$, there is at least one choice of potential basis, $\{b_\nu\}$ such that non-local effects can be *simulated* by an appropriate choice of local potential. Even for $\{b_\nu\}$ not so pathologically chosen, we might expect to see some ability to simulate non-local parts of $\hat{K}$ with a sufficiently large number of local basis functions.

A truly degenerate case

It is certainly possible that, in a finite wavefunction basis, the optimal energies given by both the OEP and HF methods coincide. One very contrived example is a basis consisting of n_e functions $a_1,\ldots,a_{n_e}$ where the $a_i(x)$ are the exact orbitals of the Hartree-Fock energy. In such a basis, no

matter what v_X may be, the P corresponding to the eigenvectors of $\hat{H}^{KS}$ is the identity matrix and the solution coincides with the Hartree-Fock solution. Without much effort one can select a basis, which ensures that the Kohn-Sham solutions are arbitrarily close to the Hartree-Fock solutions, *independent* of the choice of basis for v_X. Our thesis is that this example is generic. If the wavefunction basis is sufficiently small, and spans the HF solution, then the xOEP solution will be forced to resemble the HF solution. Staroverov et al, in contrast, have considered the wavefunction basis fixed and reasonably rich, while having too large a potential function basis.

The origin of the pathology

One problem in the identification of the source of these apparent pathologies is that it is easy to carry over intuitions from previous methods for DFT that do not apply to OEP calculations. One of these is that the energy is variational in the wavefunction basis, i.e. as one adds more elements to the basis, the optimal E value must decrease. This is true for traditional DFT and HF, but is not true for xOEP. In xOEP, $E[P]$ is variational in v_X and, hence, the local potential basis, but not the wavefunction basis, which serves merely to accurately calculate E as a functional of v_X. Additional basis functions allow for more accurate $\hat{H}^{KS}$ wavefunctions, of course, but this need not be accompanied by a decrease in $E[P]$, it may increase, in fact. For example, consider the case above in which the basis initially consists of exact HF orbitals. If the wavefunction basis is extended by a generic basis function, $E[P]$ can be expected to increase.

In summary, the origin of the apparent pathologies discussed above is that it is assumed that the bases uses to represent the potential and the wavefunctions can be varied independently, while in fact, the optimized effective potential is properly defined in terms of the exact solution of the Kohn-Sham equations for a given variational potential. Practical calculations require the use of a finite basis for the wavefunctions, which is not a problem as long as the chosen finite basis gives a sufficiently accurate representation of the exact wavefunctions. The following procedure should obtain the correct OEP solution using finite bases. (1) Choose a $\{b_\nu\}$ potential basis and an initial $\{a_\mu\}$ wavefunction basis, and solve the finite dimensional OEP problem. (2) Extend the $\{a_\mu\}$ basis with new elements, keeping $\{b_\nu\}$ fixed, until $E[P]$ converges. One then has an upper bound on the true OEP energy, variational in v_X. Augment $\{b_\nu\}$ and iterate beginning with step (1) until the energy is converged.

Plane-wave calculations

Results obtained using plane-wave bases typically converge smoothly and steadily with the plane-wave cutoff, and therefore, it may be possible to circumvent the double loop of convergence tests described above. Consider the plane-wave bases, $a_\mu(x) = e^{i\mu\dot{x}}$ and $b_\nu(x) = e^{i\nu\dot{x}}$ where $||\mu||^2 \leq \varepsilon_C$. Clearly, the eigenvectors of $\hat{H}^{KS}$ have no dependence of any potential basis elements with $||\nu||^2 > 4\varepsilon_C$. In periodic systems, the number of basis elements for potentials can be no larger that eight times the number of basis elements for wavefunctions. If the wavefunctions have n degrees of

freedom, then the potentials should have, at most, $\propto n$ degrees of freedom. The construction of Staroverov requires $\propto n^2$ basis functions for potentials and does not apply to plane wave bases. These considerations raise some doubt regarding the nature of their manipulations.

There is some reasonably physical reasoning to suspect that the dimension of potential space should be proportional (if not less than or equal) to the dimension of wavefunction space. One interpretation of the term *local* as applied to a finite basis representation of v_X is that v_X approximately commute with the position operator. In a finite wavefunction basis, the position operators are matrices of order $|\{a_\mu\}|$. Hence, the space of operators which commute with the finite basis position operators can be no more than $|\{a_\mu\}|$ dimensional.

For our own plane-wave work, we take $||v||^2 \leq \varepsilon_C$, ensuring an equal number of degrees of freedom. This also has the effect of removing certain null-vectors of the v_X minimization which ultimately arise from the ability of the wavefunction basis to represent the first-order orbital shifts used to compute the OEP gradient.

Discussion

We have found and verified an expression for the gradient of the Kohn-Sham energy with respect to the local potential appearing in the Kohn-Sham Hamiltonian. Our derivation based on the density matrix naturally provides a result that is valid at finite temperature. The cost of evaluating the optimized effective potential using this approach should be comparable to the cost of a traditional density functional calculation using standard functionals such as the LDA or GGA, but a greatly extended family of exchange-correlation functionals that have an explicit dependence on the Kohn-Sham orbitals can be considered. We have identified the source of an apparent pathology in iterative OEP calculations using a finite basis, and described a systematic procedure to obtain the correct OEP solution.

Chapter 3

Implementation of the OEP in Socorro

Our iterative approach to solving for the OEP was implemented in the Socorro electronic structure software in the orbital representation using a plane wave basis set and norm-conserving pseudopotentials. There have been no previous implementations of the OEP which address both our scientific goals (finite temperature simulations, calculation of variational quantities at the optimum) and our computational goals (production quality with pseudo-potentials, proper handling of cutoffs, and good convergence and preconditioning). We have made substantial progress towards our goals in both these areas. In the process, we have isolated and circumvented a number of computational pitfalls, which we believe have not been noted before. This section reviews our OEP implementation and discusses the computational issues that we have identified.

Review of the OEP

We have a "bare" Hamiltonian $H_0 = K + V_I$ representing the sum of kinetic energy and ionic (or otherwise external) potentials. The V_I operator is theoretically local, but in practice, it is much more computationally effective to use a non-local pseudo-potential. We will hence consider V_I to be a general operator.

We consider two energy functions

$$\begin{aligned} E_* &= H_0 \bullet \rho + V \bullet \rho + \frac{1}{\beta}(\rho \bullet \log\rho + (I-\rho) \bullet \log(I-\rho)) \\ E &= H_0 \bullet \rho + E_{HXC}(\rho) + \frac{1}{\beta}(\rho \bullet \log\rho + (I-\rho) \bullet \log(I-\rho)) \\ &= E_* + E_{HXC}(\rho) - V \bullet \rho \end{aligned}$$

where we adopt the convention $A \bullet B = \mathrm{tr}\{A^H B\}$. In the OEP formulation, we consider ρ to be a function of V defined by the condition that $I \bullet \rho = n$ and $E_*(\rho)$ is minimized with V fixed. This is equivalent to the condition that $\frac{\partial E_*}{\partial \rho} = \mu I$, for some Lagrange multiplier μ. Thus we may equivalently take $\rho(V) = (I + e^{\beta(H_0+V-\mu I)})^{-1}$, where μ is selected to ensure $I \bullet \rho(V) = n$.

We then minimize $E(\rho(V))$ as a function of V. In translating derivatives in ρ to derivatives in V

we make use of a *chain rule* for functions of Hermitian matrices,

$$\frac{\partial F(f(X))}{\partial X} = \frac{\Delta f}{\Delta x}\left[\left.\frac{\partial F(Y)}{\partial Y}\right|_{Y=f(X)}\right]$$

where the action of $\frac{\Delta f}{\Delta x}[\cdot]$, in an X diagonal basis (letting $x_i = x_{ii}$ be the eigenvalues of X), is given by

$$\left(\frac{\Delta f}{\Delta x}[Z]\right)_{ij} = \frac{f(x_i) - f(x_j)}{x_i - x_j} Z_{ij}$$

where $\frac{f(x_i)-f(x_j)}{x_i-x_j}$ is a divided difference (i.e. taking derivatives when $x_i = x_j$).

For the application of interest, our function is $\omega(\varepsilon) = \frac{1}{1+e^{\beta\varepsilon}}$. The gradient of E as a function of V is

$$\begin{aligned}\frac{\partial E}{\partial V} &= \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E}{\partial \rho} + \nu I\right] \\ &= \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E_{HXC}}{\partial \rho} - V + \nu I\right].\end{aligned}$$

where ν (actually the derivative of μ) is selected so that $I \bullet \frac{\partial E}{\partial V} = 0$. Thus, for OEP, where V is restricted to be a local operator, and the gradient is

$$\frac{\partial E}{\partial V} = \mathrm{diag}\left[\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E_{HXC}}{\partial \rho} - V + \nu I\right]\right],$$

where we understand the diag above to be the diagonal piece of a Hermitian matrix taken in real space.

At this point, we should comment that when $\frac{\partial E_{HXC}}{\partial \rho}$ is a local operator, as in the case of LDA, then one can motivate the so-called *self-consistent* iteration $V \leftarrow \frac{\partial E_{HXC}}{\partial \rho} + \nu I$, which should converge nicely as long as $\frac{\Delta\omega}{\Delta\varepsilon}$ does not change much. This is one reason why LDA calculations have never needed to think about $\frac{\Delta\omega}{\Delta\varepsilon}$ or a gradient formulation of LDA.

As an aside, in Hartree-Fock it is the case that $\frac{\partial E_{HXC}}{\partial \rho}$ is non-local but that there is no restriction on ρ, which can be thought (perversely) as there being no locality restriction on V, and thus HF is effectively doing $V \leftarrow \frac{\partial E_{HXC}}{\partial \rho} + \nu I$ as well.

However, when doing exact exchange $\frac{\partial E_{HXC}}{\partial \rho}$ is a non-local operator, and it is not clear how a fixed-point scheme, similar to that of LDA could be generated. This leaves us with the gradient search as our best available approach.

Review of the application of $\frac{\Delta\omega}{\Delta\varepsilon}[\cdot]$

A full diagonalization of $H = H_0 + V$ is computationally impractical for large basis sets. We review here how the OEP gradient can be found by solving for $\frac{\Delta\omega}{\Delta\varepsilon}[Z]$ (for Z a Hermitian operators) iteratively on a set of orbitals $\phi = \begin{pmatrix}\phi_1 & \cdots & \phi_N\end{pmatrix}$ which are the lowest N eigenvectors of H with eigenvalues ε_i and occupations ω_i (we assuming $\omega_i \sim 0$ for $i > N$).

We will generalize slightly, to the computation of $\frac{\Delta\omega}{\Delta\varepsilon}[Z + \nu I]$, where ν is chosen so that $\mathrm{tr}\{\frac{\Delta\omega}{\Delta\varepsilon}[Z + \nu I]\} = t$ (in the case where there is no trace constraint, $\nu = 0$. Let $\bar{E}$ and $\bar{J}$ be $N \times N$ matrices given by

$$\begin{aligned}\bar{E} &= \phi^\dagger Z \phi \\ \bar{J} &= \phi^\dagger \frac{\Delta\omega}{\Delta\varepsilon}[Z + \nu I]\,\phi = \left[\frac{\omega_i - \omega_j}{\varepsilon_i - \varepsilon_j}\right] \circ (\bar{E} + \nu I),\end{aligned}$$

where $\left[\frac{\omega_i - \omega_j}{\varepsilon_i - \varepsilon_j}\right]$ is the $N \times N$ matrix of divided differences, $\circ$ denotes elementwise multiplication and ν is chosen such that $\mathrm{tr}\{\bar{J}\} = t$. Let $\psi_\perp$ satisfy $\phi^\dagger \psi_\perp = 0$ and

$$H\psi_\perp - \psi_\perp \Lambda = -(I - \phi\phi^\dagger) Z \phi \Omega, \tag{3.1}$$

where Λ, Ω are diagonal matrices with entries ε_i, ω_i respectively. Then

$$\frac{\Delta\omega}{\Delta\varepsilon}[Z + \nu I] = \psi_\perp + \frac{1}{2}\phi\bar{J}.$$

We see that we can compute $\frac{\Delta\omega}{\Delta\varepsilon}[Z + \nu I]$ from products of the form $Z\phi$. However, this does require an iterative solution to (3.1). It is customary to allow fairly loose tolerances for the underlying eigenproblem when far from convergence. Thus, ϕ may be substantially different from the true lowest eigenvalues of H and $[\phi\phi^\dagger, H]$ may not be small. In that case, we can do additional projection,

$$(I - \phi\phi^\dagger) H \psi_\perp - \psi_\perp \Lambda = -(I - \phi\phi^\dagger) Z \phi \Omega,$$

to ensure that we take a principle submatrix of the $[H, \cdot]$ operator on the left hand side, and thus a well-posed problem. Even so, if an iterative solution method requiring positive definiteness is used, poorly converged ϕ can lead to a principle submatrix which is not positive definite.

Generally, we have found that we require our tolerances for the eigenvectors, ϕ, to be a bit higher than those required for self-consistent LDA.

Iterative Algorithms

There are two main iterative loops involved in the iterative OEP algorithm: an inner loop that solves a linear system in order to evaluate the gradient at a given potential and an outer loop that uses the resulting gradient to minimize the energy with respect to the potential.

The conjugate gradient algorithm was used to solve the linear systems in the inner loop. Standard preconditioning techniques identical to those used in iteratively solving for the eigenvectors of the Hamiltonian in a standard plane wave DFT code are effective in this case. Using the conjugate algorithm, the computational cost of solving the set of linear systems determining $\psi_\perp$ is comparable to the cost of solving the Kohn-Sham eigenproblem for ϕ. Therefore, each gradient evaluation is approximately as computationally expensive as one step of the self-consistency loop in a standard DFT code.

The outer loop, in which the energy is minimized, replaces the self-consistency iteration in a standard DFT code. Minimization requires different algorithms than self-consistency, and therefore, a considerable amount of effort was spent optimizing the outer iteration. Since line minimization is expensive and the Hessian in this case is difficult to apply, we converged on a *fixed step* approach such as Richardson iteration.

The basic relaxation step is of the form

$$x_{i+1} = x_i - \gamma g_i,$$

where $g_i = \nabla f(x_i)$, which linearizes to

$$x_{i+1} = x_i + \gamma(b - Ax_i).$$

The iteration on the error $e_i = x_i - x_*$ is given by

$$\begin{aligned} e_{i+1} &= (I - \gamma A)e_i \\ e_i &= (I - \gamma A)^i e_0 \end{aligned}$$

from which we see that $e_i \to 0$ only if $-1 \leq 1 - \gamma\lambda \leq 1$ for all eigenvalues λ of A.

If the eigenvalues of A occur in $(0,1)$ then convergence occurs when $\gamma \leq 2$ and $\gamma = 2$ is the largest convergent step size. In practice, we do not start with the spectrum of A in $(0,1)$, but we scale all gradients by an empirically determined factor to place the eigenvalues of A between $(0,1)$ in a more-or-less centered fashion.

With $\gamma = 2$, if $\lambda_{\min} = \varepsilon$ or $\lambda_{\max} = 1 - \varepsilon$, the error in the extreme modes will be $|1 - 2\varepsilon|^i$. Thus convergence is eventually dominated by the value of $\kappa = \frac{1}{\varepsilon}$, which is approximately equal to the condition number of A.

One technique which works well when good estimates of the Hessian norm and condition number are available is *Chebyshev acceleration*. One considers a Chebyshev polynomial, T_i, affinely translated so that $(-1,1) \rightarrow (\varepsilon,1)$ where the eigenvalues of the Hessian occur in $(\varepsilon,1)$. The goal is to have $e_i = \frac{1}{T_i(\beta)} T_i(\beta - \alpha A) e_0$ where $\alpha = \frac{2}{1-\varepsilon}$ and $\beta = \frac{1+\varepsilon}{1-\varepsilon}$. The Chebyshev polynomials, given by

$$\begin{aligned} T_0(y) &= 1 \\ T_1(y) &= y \\ T_{i+1}(y) &= 2yT_i(y) - T_{i-1}(y). \end{aligned}$$

The error will decrease roughly as $\frac{1}{T_i(\beta)} \sim \left(\frac{1-\sqrt{\varepsilon}}{1+\sqrt{\varepsilon}}\right)^{-i}$ with a need for periodic resets to correct non-linearities (which should be reasonable to perform after $2/\sqrt{\varepsilon}$ steps).

The recurrences for Chebyshev polynomials can be re-arranged a bit to give

$$\begin{aligned} \tau_0 &= 1 \\ \tau_1 &= 1 \\ \tau_{i+1} &= 2\tau_i - \beta\tau_{i-1} \\ x_1 &= x_0 - \alpha g_0 \\ x_{i+1} &= x_i + \frac{1}{\tau_{i+1}} \left(\tau_{i-1}\beta(x_i - x_{i-1}) - 2\tau_i \alpha g_i\right) \end{aligned}$$

and thus we need only keep around a *velocity* vector $x_i - x_{i-1}$ and mix it with the gradient to obtain the update.

With a good preconditioner in place (such as that discussed in the next section), it is reasonable to assume that Hessian norms and condition numbers are fairly insensitive to problem size. We have observed that the same pair of parameters achieve nearly the same convergence on 2 atom, 8 atom, and 64 atom silicon, supporting this intuition. This also suggests that one might use smaller systems to estimate the optimal convergence parameters for larger ones.

Preconditioning the Outer Loop

Effient convergence of the outer loop in the iterative OEP algorithm requires a good preconditioner for the minimization algorithm. We have developed a preconditioner based on the OEP gradient in the non-interacting free metal

Since the system is non-interacting, the E_{HXC} term vanishes and $V = 0$ is optimal. The gradient is

$$\frac{\partial E}{\partial V} = \mathrm{diag}\left(\frac{\Delta\omega}{\Delta\varepsilon}[0]\right)$$

which can be varied by $V \to 0 + \delta V$ to obtain the variation of the gradient,

$$\begin{aligned}\delta\frac{\partial E}{\partial V} &= \operatorname{diag}\left(\frac{\Delta\omega}{\Delta\varepsilon}[-\delta V]\right) + \operatorname{diag}\left(\left(\delta\frac{\Delta\omega}{\Delta\varepsilon}\right)[0]\right) \\ &= -\operatorname{diag}\left(\frac{\Delta\omega}{\Delta\varepsilon}[\delta V]\right)\end{aligned}$$

Thus we see that the Hessian is given by the action of the Jacobian restricted to local operators δV. There being no possibility of confusion, we will drop the δ and use V for a variation of the (vanishing) optimum effective potential.

Let us suppose that we have a discrete set of wave numbers $K \subset \mathbb{R}^3$, which we will use to represent the occupied orbitals of an idealized system. In this case, the density operator is

$$\rho(\bar{x},\bar{x}') = \sum_{\bar{k}\in K} |\bar{k}><\bar{k}| = \frac{1}{N}\sum_{\bar{k}\in K} e^{-i\bar{k}\cdot(\bar{x}-\bar{x}')}.$$

with $\rho(\bar{x},\bar{x}) = n_e/N$ where $n_e = |K|$ is the number of electrons and N is the unit volume.

We can carry out a general derivation for planewaves in finite temperature, by taking $D_{\bar{k}\bar{k}'}(\beta) = \frac{\omega(\varepsilon_{\bar{k}})-\omega(\varepsilon_{\bar{k}'})}{\varepsilon_{\bar{k}}-\varepsilon_{\bar{k}'}}$ and thus the localized Jacobian is given by

$$\begin{aligned}V(\bar{x}) &\to \sum_{\bar{k},\bar{k}'} D_{\bar{k}\bar{k}'}(\beta)|\bar{k}><\bar{k}|V|\bar{k}'><\bar{k}'| \\ &= \sum_{\bar{k}\bar{k}'} e^{-i\bar{k}\cdot\bar{x}}e^{i\bar{k}'\cdot\bar{x}}D_{\bar{k}\bar{k}'}(\beta)\int d^3\bar{x}'e^{i\bar{k}\cdot\bar{x}'}e^{-i\bar{k}'\cdot\bar{x}'}V(\bar{x}') \\ &= \sum_{\bar{k}\bar{k}'} D_{\bar{k}\bar{k}'}(\beta)\int d^3\bar{x}'e^{-i(\bar{k}-\bar{k}')\cdot(\bar{x}-\bar{x}')}V(\bar{x}') \\ &= \sum_{\bar{\alpha}} d_{\bar{\alpha}}(\beta)\int d^3\bar{x}'e^{-i\bar{\alpha}\cdot(\bar{x}-\bar{x}')}V(\bar{x}') \\ &= \sum_{\bar{\alpha}} d_{\bar{\alpha}}(\beta)|\bar{\alpha}><\bar{\alpha}|V\end{aligned}$$

where $d_{\bar{\alpha}}(\beta) = \sum_{\bar{k}-\bar{k}'=\bar{\alpha}} D_{\bar{k}\bar{k}'}(\beta)$. We notes that if $\int d^3\bar{x}V(\bar{x})$ vanishes, then so does the integral of the RHS, thus $v = 0$, and the sum over $\bar{\alpha}$ can be restricted to $\bar{\alpha} \neq \bar{0}$.

We may consider a more specific metallic case where the possible states are all $\mathbb{R}^3$ with

$$\begin{aligned}\varepsilon_{\bar{k}} &= ||\bar{k}||^2 \\ \omega(\bar{k}) &= \frac{1}{1+e^{\beta(||\bar{k}||^2-\mu)}}\end{aligned}$$

where μ is selected by $\int \frac{4\pi k^2 dk}{1+e^{\beta(k^2-\mu)}} = n$, for some constant n.

Passing from sums to integrals we find (taking $\alpha = ||\bar{\alpha}||$),

$$\begin{aligned}
d_\alpha(\beta) &= \sum_{\bar{k}} D_{(\bar{k}+\frac{1}{2}\bar{\alpha})(\bar{k}-\frac{1}{2}\bar{\alpha})} \\
&= \int \frac{\omega(\bar{k}+\frac{1}{2}\bar{\alpha}) - \omega(\bar{k}-\frac{1}{2}\bar{\alpha})}{2k\cdot\bar{\alpha}} d^3\bar{k} \\
d_\alpha(\beta) &= \int \frac{\frac{1}{1+e^{\beta((x+\frac{1}{2}\alpha)^2+y^2+z^2-\mu)}} - \frac{1}{1+e^{\beta((x-\frac{1}{2}\alpha)^2+y^2+z^2-\mu)}}}{2x\alpha} dxdydz \\
&= \int_{r\in\mathbb{R}_+}\int_{x\in\mathbb{R}} \pi r \frac{\frac{1}{1+e^{\beta((x+\frac{1}{2}\alpha)^2+r^2-\mu)}} - \frac{1}{1+e^{\beta((x-\frac{1}{2}\alpha)^2+r^2-\mu)}}}{x\alpha} dxdr \\
&= \int_{x\in\mathbb{R}} \pi \frac{\log\frac{1+e^{\beta((x-\frac{1}{2}\alpha)^2+r^2-\mu)}}{1+e^{\beta((x+\frac{1}{2}\alpha)^2+r^2-\mu)}}\Big|_{r=0}^{\infty}}{2\beta x\alpha} dx \\
&= \int_{x\in\mathbb{R}} \pi \frac{\log\frac{e^{\beta((x-\frac{1}{2}\alpha)^2-\mu)}}{e^{\beta((x+\frac{1}{2}\alpha)^2-\mu)}} - \log\frac{1+e^{\beta((x-\frac{1}{2}\alpha)^2-\mu)}}{1+e^{\beta((x+\frac{1}{2}\alpha)^2-\mu)}}}{2\beta x\alpha} dx \\
&= \int_{x\in\mathbb{R}} \frac{\pi}{2x\alpha}\frac{1}{\beta}\log\frac{e^{-\beta((x+\frac{1}{2}\alpha)^2-\mu)}+1}{e^{-\beta((x-\frac{1}{2}\alpha)^2-\mu)}+1} dx.
\end{aligned}$$

Taking the limit as $\beta\to\infty$, noting that $\lim_{\beta\to\infty}\frac{1}{\beta}\log(e^{\beta c}+1) = \max\{c,0\}$ for all c,

$$\begin{aligned}
d_\alpha(\infty) &= \int_{x\in\mathbb{R}} \frac{\pi}{2x\alpha}\left(\max\left\{\mu-(x+\frac{1}{2}\alpha)^2,0\right\} - \max\left\{\mu-(x-\frac{1}{2}\alpha)^2,0\right\}\right) dx \\
&= \int_{x=\frac{1}{2}\alpha-\sqrt{\mu}}^{\frac{1}{2}\alpha+\sqrt{\mu}} \frac{\pi}{x\alpha}\left((x-\frac{1}{2}\alpha)^2-\mu\right) dx \\
&= \int_{x=\frac{1}{2}\alpha-\sqrt{\mu}}^{\frac{1}{2}\alpha+\sqrt{\mu}} \pi\left(\frac{x}{\alpha}-1+\frac{\frac{1}{4}\alpha^2-\mu}{x\alpha}\right) dx \\
&= \pi\left(\frac{x^2}{2\alpha}-x+\frac{\frac{1}{4}\alpha^2-\mu}{\alpha}\log|x|\right)\Bigg|_{x=\frac{1}{2}\alpha-\sqrt{\mu}}^{\frac{1}{2}\alpha+\sqrt{\mu}} \\
&= -\pi\left(\sqrt{\mu}-\frac{\frac{1}{4}\alpha^2-\mu}{\alpha}\log\frac{\frac{1}{2}\alpha+\sqrt{\mu}}{|\frac{1}{2}\alpha-\sqrt{\mu}|}\right).
\end{aligned}$$

We now note some of the features of $d_\alpha(\infty)$. It is continuous everywhere, taking the value $-\pi\sqrt{\mu}$ at $\alpha = 2\sqrt{\mu}$. It is differentiable for all $\alpha \neq 2\sqrt{\mu}$ (where it is infinite), which is associated with the Friedel oscillations in metals.

For large α,

$$\begin{aligned} d_\alpha(\infty) &= -\pi\left(\sqrt{\mu} - \frac{\frac{1}{4}\alpha^2-\mu}{\alpha}\log\frac{1+\frac{2\sqrt{\mu}}{\alpha}}{1-\frac{2\sqrt{\mu}}{\alpha}}\right) \\ &= -\pi\left(\sqrt{\mu} - \left(\frac{1}{4}\alpha-\frac{\mu}{\alpha}\right)2\left(\frac{2\sqrt{\mu}}{\alpha}+\frac{1}{3}\left(\frac{2\sqrt{\mu}}{\alpha}\right)^3+\cdots\right)\right) \\ &= -\pi\left(\frac{4\mu^{3/2}}{\alpha^2} - \left(\frac{1}{2}\alpha-\frac{2\mu}{\alpha}\right)\left(\frac{1}{3}\left(\frac{2\sqrt{\mu}}{\alpha}\right)^3+\cdots\right)\right) \\ &\sim -\pi\frac{8\mu^{3/2}}{3\alpha^2}. \end{aligned}$$

For small α,

$$\begin{aligned} d_\alpha(\infty) &= -\pi\left(\sqrt{\mu} - \frac{\frac{1}{4}\alpha^2-\mu}{\alpha}\log\frac{1+\frac{\alpha}{2\sqrt{\mu}}}{1-\frac{\alpha}{2}\sqrt{\mu}}\right) \\ &= -\pi\left(\sqrt{\mu} - \left(\frac{1}{4}\alpha-\frac{\mu}{\alpha}\right)2\left(\frac{\alpha}{2\sqrt{\mu}}+\frac{1}{3}\left(\frac{\alpha}{2\sqrt{\mu}}\right)^3+\cdots\right)\right) \\ &\sim -\pi\left(2\sqrt{\mu} - \frac{\alpha^2}{6\sqrt{\mu}}\right). \end{aligned}$$

A bit of empirical curve fitting and rearrangement of expressions then shows that

$$C\sqrt{\mu} \le d_\alpha(\infty)\cdot\left(\frac{1}{1+\frac{\alpha^2}{\mu}}+\frac{3\alpha^2}{4\mu}\right) \le 2C\sqrt{\mu}$$

for some C, which implies that

$$\frac{1}{1-\frac{\nabla^2}{\mu}} - \frac{3\nabla^2}{4\mu} \quad \text{or} \quad 1-\frac{3}{4\mu}\nabla^2$$

should be a good preconditioner in the low temperature regime. This is the preconditioner that we use in our OEP implementation, and it seems to be doing a good job even on non-metallic systems such as silicon and at substantial finite temperatures.

Potential cut-offs

Another issue that arises in the implementation of the iterative OEP algorithm is making sure that we do not include modes in the variational space of the potential that have an extremely weak effect on the energy evaluated using a given plane wave representation of the waevefunctions. Such modes lead to an extremely poorly conditioned minimization problem and very poor convergence. This issue is closely related to the problem discussed above in the theory section of making sure that the Kohn-Sham equations are accurately solved for each potential occuring during the minimization of the potential.

Let us consider the case where we have used a planewave basis with some cutoff $||\bar{k}|| < k_W$ and fully diagonalized $H = H_0 + V$ into a complete eigenbasis $\phi_1, \ldots, \phi_N$ (where N is the dimension of the planewave basis). Let us also assume that we have likewise represented the local operator, V, in a planewave basis with a cutoff of k_V.

We consider the case where $V = V_* + \varepsilon e^{i\bar{\alpha}\cdot\bar{x}}$, where V_* is the optimal and v is small. Then the gradient at V is

$$\frac{\partial E}{\partial V} = \mathrm{diag}\left(\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E_{HXC}(\rho)}{\partial\rho} - V_* - \varepsilon e^{i\bar{\alpha}\cdot\bar{x}} + (v_* + v\gamma)I\right]\right) = -\varepsilon\,\mathrm{diag}\left(\frac{\Delta\omega}{\Delta\varepsilon}\left[e^{i\bar{\alpha}\cdot\bar{x}} - \gamma I\right]\right)$$

where γ is chosen to make $\frac{\partial E}{\partial V}$ traceless. In terms of the complete eigenbasis,

$$\frac{\Delta\omega}{\Delta\varepsilon}\left[e^{i\bar{\alpha}\cdot\bar{x}}\right] = -\sum_{ij}\left(\frac{\omega_i - \omega_j}{\varepsilon_i - \varepsilon_j}\right)\left(\phi_i^\dagger e^{i\bar{\alpha}\cdot\bar{x}}\phi_j\right)\phi_i\phi_j^\dagger, \tag{3.2}$$

from which we see a vanishing gradient for $||\bar{\alpha}|| > 2k_W$. This is to be expected, however, since the coupling between V and ρ in E_* is $V \bullet \rho$ and $e^{i\bar{\alpha}\cdot\bar{x}} \bullet \rho$ identically vanishes when $||\bar{\alpha}|| > 2k_W$. One way to look at this is that high frequency modes of V have no effect on ρ in the presence of cutoffs, and thus, no effect on the energy E. If we started the minimization from $V = V_* + \varepsilon e^{i\bar{\alpha}\cdot\bar{x}}$, we could not expect to see ε decrease.

However, this is not enough. We have observed that with V cutoff at $2k_W$ the convergence of the gradient search is very sub-linear.

In fact, any first order perturbation to V_* which gives a second order gradient could not be expected to decrease rapidly (i.e. linearly) in a gradient-based minimization. Such perturbations are then null vectors of the objective function $E(\rho(V))$.

Consider that in order for a term in (3.2) to contribute significantly to the sum, at least one of ϕ_i or ϕ_j must have ω_i, ω_j non-vanishing, i.e. be *partially occupied*. In a typical calculation, the partially occupied states have planewave components which become vanishingly small for $||\bar{k}|| > k_0$ with k_0 independent of k_W so long as k_W is picked to be well enough above k_0 (an extreme case is the

free non-interacting metal in which k_0 is minimal and independent of k_W for $k_W > k_0$). Thus if $||\bar{\alpha}|| > k_W + k_0$ the first order contribution to ρ will be likewise small and ill-conditioning of the optimization results.

This ill-conditioning is entirely non-physical, coming from the choice of k_W. If we consider the effect of increasing k_W, k_0 would change negligibly, and thus $k_W + k_0$ would increase and more $\bar{\alpha}$'s would be able to significantly contribute (3.2). In non-OEP formulations, the criterion for k_W is to provide a good representation of the partially occupied orbitals. However, the OEP gradient couples unoccupied and partially occupied orbitals, and this coupling can only be represented faithfully so long as the unoccupied orbitals can be represented faithfully.

We have used a more stringent cutoff for V therefore, taking it to be k_W instead of the customary $2k_W$. We have found, with experiments on silicon, that the number of correctly converged digits in the resulting energies did not change. This is to be expected, since a tiny contribution to the gradient indicates a tiny contribution to the overall energy.

OEP Hellman-Feynman correction

Once an OEP calculation has been performed, it is important to be able to calculate variational quantities (such as forces) at the optimum. One of the surprising results we've obtained is the failure of the Hellman-Feynman theorem in general OEP problems.

The Hellman-Feynman theorem is the basis for doing perturbative analysis of LDA approximations or Hartree-Fock to obtain variations in the ground state energy as a result of a varying Hamiltonian. It says that if the bare Hamiltonian is a linear functon $H_0(s) = H_0 + sH_1$, and $E_{min}(s)$ is the minimum of E (as a function of V) with the given $H_0(s)$, then

$$\frac{dE_{min}}{ds} = H_1 \bullet \rho(V)$$

where V is the minimizer at $s = 0$, giving a linearization of the energy which is independent of any first derivatives of V at the minimum. These first order variations of the minimum energy are the basis for the calculation of a variety of bulk material properties, like conductivity, the dielectric constant, as well as stress and strain. This is what one would generally expect from an arbitrary function of the form $f(s) = \min_x F(s,x)$, a short sketch of why being,

$$\begin{aligned} \frac{\partial}{\partial x}F(x,s) &= 0 \\ \frac{d}{ds}F(x(s),s) &= \frac{\partial}{\partial x}F(x(s),s)\frac{dx(s)}{ds} + \frac{\partial}{\partial s}F(x(s),s) \\ &= 0 + \frac{\partial}{\partial s}F(x(s),s). \end{aligned}$$

However, our problem is analogous to

$$\begin{aligned}
\frac{\partial}{\partial y}F(x(y,s),s) &= \frac{\partial}{\partial x}F(x(y,s),s)\frac{\partial x}{\partial y} = 0 \\
\frac{d}{ds}F(x(y(s),s),s) &= \frac{\partial}{\partial x}F(x(y,s),s)\left(\frac{\partial x}{\partial y}\frac{dy(s)}{ds}+\frac{\partial x}{\partial s}\right)+\frac{\partial}{\partial s}F(x(y(s),s),s) \\
&= \frac{\partial}{\partial x}F(x(y,s),s)\frac{\partial x}{\partial s}+\frac{\partial}{\partial s}F(x(y(s),s),s),
\end{aligned}$$

and it is this $\frac{\partial}{\partial x}F(x(y,s),s)\frac{\partial x}{\partial s}$ term, which comes from the fact that the relation between x and y is dependent on s, that we must account for.

For the OEP formulation,

$$\begin{aligned}
\frac{d}{ds}E &= \frac{\partial E}{\partial H}\bullet H_1+\frac{\partial E}{\partial s} \\
&= \left(\frac{\partial E_{HXC}}{\partial \rho}-V+\mu I\right)\bullet\frac{\Delta\omega}{\Delta\varepsilon}[H_1]+\mathrm{tr}\{\rho H_1\} \\
&= \frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E_{HXC}}{\partial \rho}-V+\mu I\right]\bullet H_1+\mathrm{tr}\{\rho H_1\}
\end{aligned}$$

The second term vanishes if

- H_1 is local (since the local part of $\frac{\Delta\omega}{\Delta\varepsilon}\left[\frac{\partial E_{HXC}}{\partial \rho}-V+\mu I\right]$ vanishes)
- LDA: $\frac{\partial E_{HXC}}{\partial \rho}$ is local (since $\frac{\partial E_{HXC}}{\partial \rho}-V+\mu I$ vanishes)
- Hartree-Fock: V is not restricted to be local (since $\frac{\partial E_{HXC}}{\partial \rho}-V+\mu I$ vanishes)

Thus, this correction is not present in the previous methods of LDA and Hartree-Fock. Corrections to higher order derivatives are thus also expected to appear in an OEP problem with non-local $\frac{\partial E_{HXC}}{\partial \rho}$.

Chapter 4

Tests and Applications

We applied our implementation of EXX using a plane-wave basis and our iterative OEP algorithm to study small molecules, bulk semiconductors, and defects in silicon. In these calculations, we neglected correlation and take E_{HXC} to be the sum of the Hartree and exchange energies.

Convergence Tests

We repeated the convergence tests for our iterative OEP algorithm described above for the LDA functional using the EXX. This verifies that the method works correctly using a E_{HXC} that can not be written as an explicit functional of the electronic density. The test system consisted of a two atom unit cell of silicon in the diamond structure. We used a 20 Rydberg plane-wave cutoff and a $2\times2\times2$ Monkhorst-Pack k-point sampling. This k-point sampling does not give a converged total energy, but this is not an issue for the purpose of testing our approach. Two electronic temperatures were used: (1) Room Temperature ($k_BT = 25.67$ meV), and (2) High Temperature ($k_BT = 1.0$ eV).

In order to test the correctness of our gradient, we applied the finite difference approach to the EXX energy functional. During each of a series of steps, the value of V at each point on a real-space grid was varied by a small ($o(10^{-4})$) random perturbation $\Delta V(r)$. During this random walk, the energy and the gradient were evaluated at each step. A linear approximation to the change in energy during each step is given by

$$\Delta E \approx \int \frac{\partial E}{\partial V(r)} \Delta V(r) dr. \tag{4.1}$$

For the small steps taken in this test, we would expect this linear approximation to be accurate if the gradient is accurate. Therefore, we can compare this predicted energy change to the actual energy change observed during the random walk. The results of this comparison for the high temperature case are shown in figure 4.1. Since the step direction is random, this represents a very stringent test of the accuracy of the gradient, and we believe that the excellent agreement between the predicted and actual energy changes demonstrates that our approach gives an accurate gradient, even at large electronic temperatures.

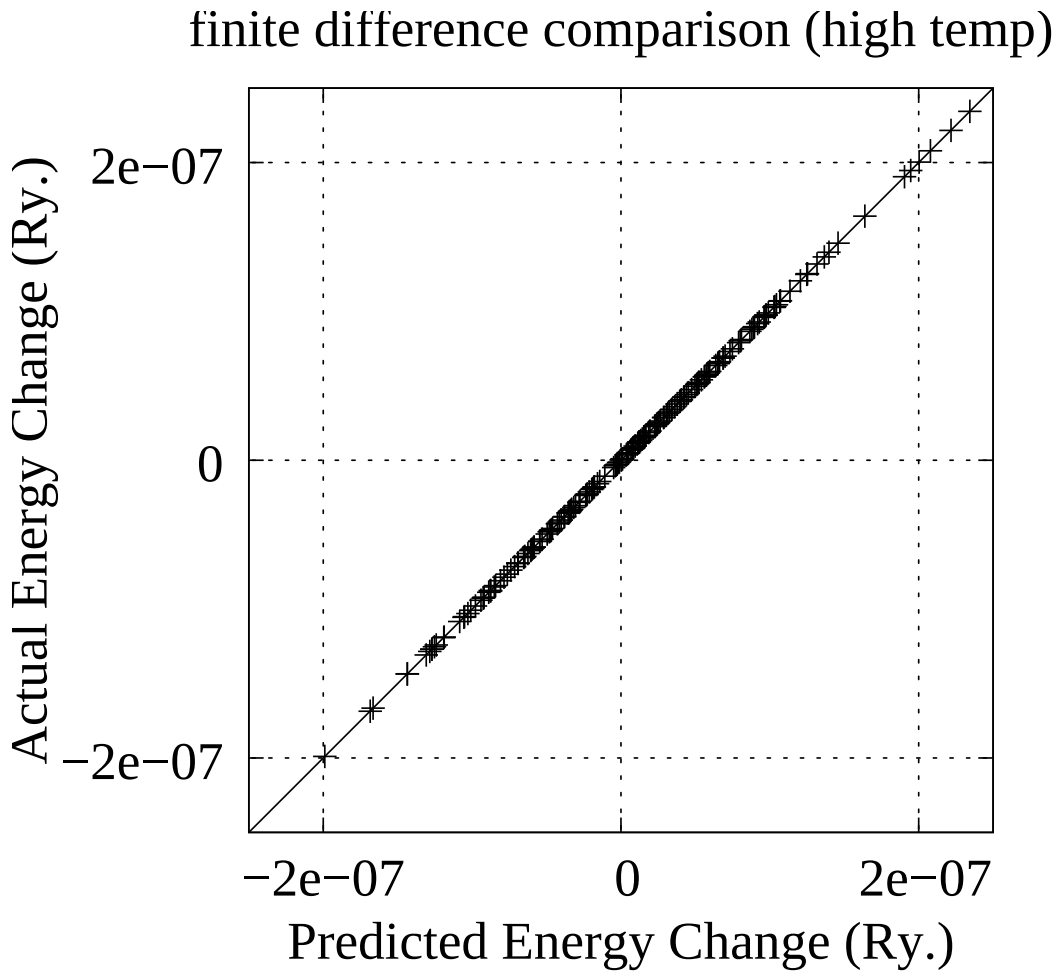


Figure 4.1. A comparison of the energy change predicted from the gradient and the actual energy change observed during a random walk in the potential. The crosses are the calculated values. The solid line is a guide to the eye representing perfect agreement.

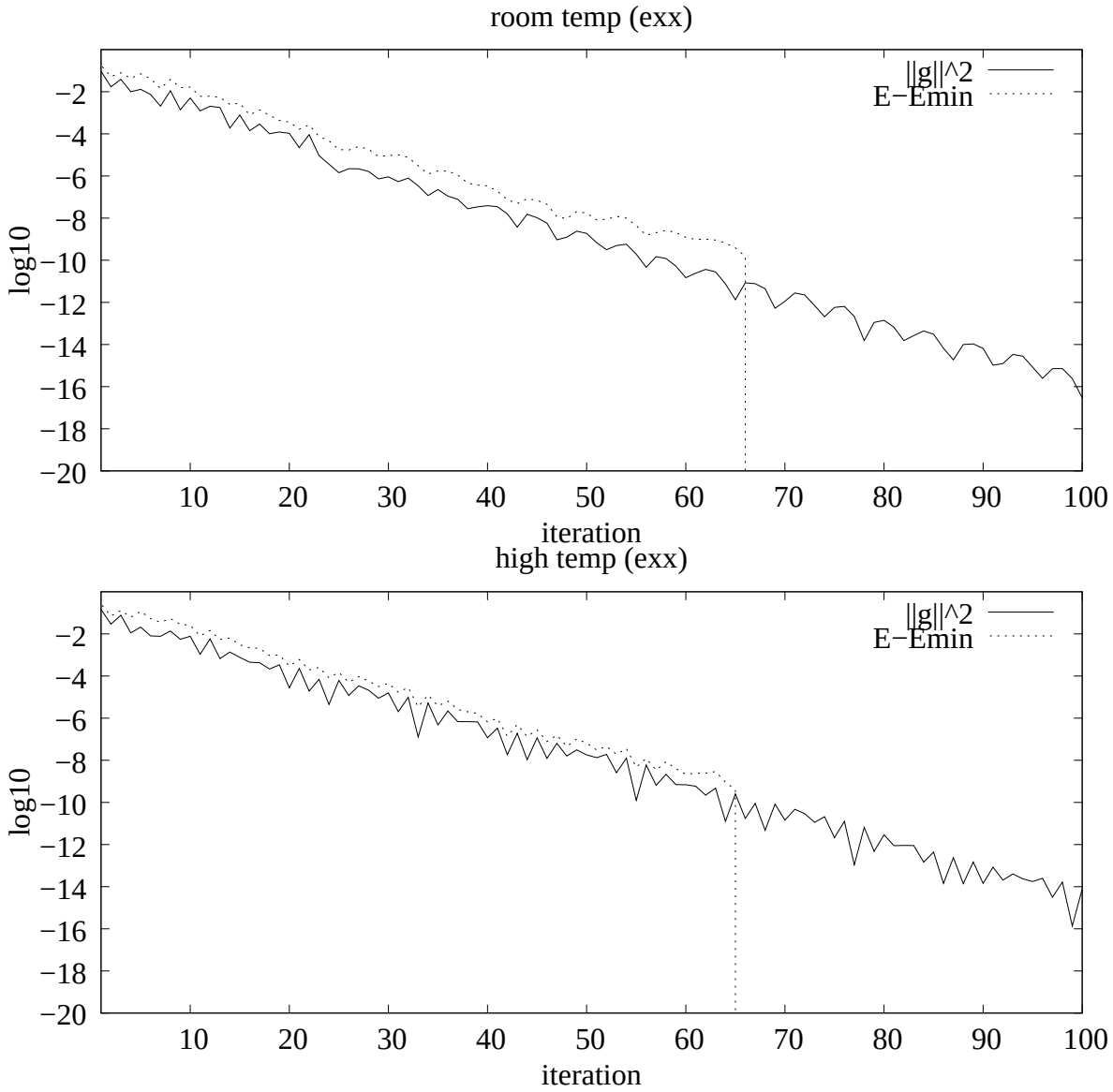


Figure 4.2. The convergence of the **exact exchange** energy, as well as the square-norm of gradient, during the iterative minimization on a log scale. The top plot was run at room temperature (25.67 meV) and the bottom at high temperature (1 eV). $g = \frac{\partial E}{\partial V(r)}$ from (2.25).

The convergence of the energy of our EXX test system during the application of the iterative OEP algorithm is shown in Fig. 4.2. The errors in the energy were approximated by comparing to the converged result. As in the LDA case, we found that the convergence of the EXX energy is only weakly dependent on the electronic temperature. The asymptotic rate of convergence obtained in the iterative OEP approach is not as rapid as the highly optimized mixing methods typically used in self-consistent calculations, but a reasonable accuracy for practical purposes (10^{-4} Ry.) can be obtained easily.

EXX results for H_2 and bulk Si and Ge

The bond length of the H_2 molecule was calculated using our plane-wave implementation of EXX. Pseudopotentials were not used, so the hydrogen nucleus was represented by a Coulomb divergence regularized by the finite plane-wave basis set used in the calculations. Convergence with respect to supercell size (which gives the distance between molecules) and plane-wave cutoff was established. The resulting bond length was 1.40 atomic units, identical to the experimental result, and in good agreement with the EXX result of 1.39 atomic units found using a Gaussian basis set and discussed elsewhere in this report. This shows that, like conventional DFT exchange-correlation functionals such as LDA and GGA, the EXX can give good structural properties.

The band gaps of Si and Ge were also calculated. Pseudopotentials generated within the KLI approximation using the FHI code were used. The results can be compared to EXX results obtained using a very different technique by Städele et al [12]. In the case of Si, our calculations gave a band gap of 1.28 eV, in good agreement with the values 1.23 eV from other EXX calculations, and 1.17 eV from experiment. In the case of Ge, our calculations gave a band gap of 0.72 eV, which can e compared to 0.94 eV from other EXX calculations, and 0.66 eV from experiment. The differences between our EXX results and those of Städele et al are within the expected uncertainties due to differences between our pseudopotentials (KLI) and their pseudopotentials (full EXX) [12]. As observed before, EXX band gap values are in much better agreement with experiment than conventional exchange-correlation functionals such as LDA and GGA. This is particularly dramatic for the case of Ge, where LDA predicts metallic behavior (the absence of a band gap).

Finally, we calculated the Si band gap as a function of temperature using EXX and LDA. The results are shown in Fig. 4.3. Note that these calculations neglect the effects of ion motion, such as thermal expansion, and therefore, they can not be straightforwardly compared to finite temperature experimental results. As discussed above, the EXX result at low temperature is much closer to the experimental result than the LDA result. The curves giving the calculated results come in flat at zero temperature. This is the expected behavior for all semiconductors when using energy functionals, such as EXX and LDA, that depend only on the temperature via the electronic density matrix. This behavior results from the absence of a linear dependence of the density matrix of a semiconductor on the temperature at zero temperature. In contrast, GW calculations show a decrease in the band gap with temperature at small temperatures. The EXX band gap is observed to increase more with temperature than the LDA result, and there are signs of a plateau in the EXX

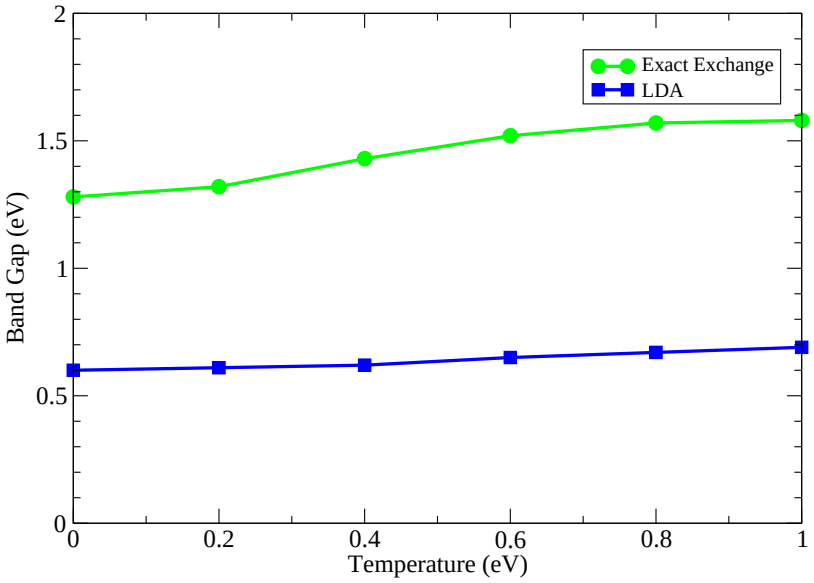


Figure 4.3. Our calculated band gap of silicon as a function of temperature using EXX and the LDA.

band gap at high temperature.

EXX results for the silicon interstitial

Due to the favorable scaling of our iterative OEP algorithm, we were able to perform unprecedented EXX calculations for defects in 216 atom unit cells of silicon. Figure 4.4 shows the results for the -2 charge state of the silicon self-interstitial in the ground state ([110] split) configuration. The EXX calculations were performed using PBE relaxed atomic positions. As mentioned before, the EXX band gap is in good agreement with experiment and much larger than the PBE band gap. The allowable charge states of the silicon self-interstitial range from +2 to -2, and therefore, the -2 charge state should have two filled levels within the band gap. This is observed in the EXX case. However, in the PBE case, the lower level has moved down and submerged into the valence band, while the upper level is unphysically close to the conduction band edge, which, in the presence of defect band dispersion, can lead to partial loss of charge from the upper level to the conduction band edge. Such a loss of charge from a defect level can lead to incorrect results for the properties of defects in conventional DFT calculations. Our results show that the EXX provides a promising approach to avoiding unphysical band overlap problems and the associated errors in defect calculations.

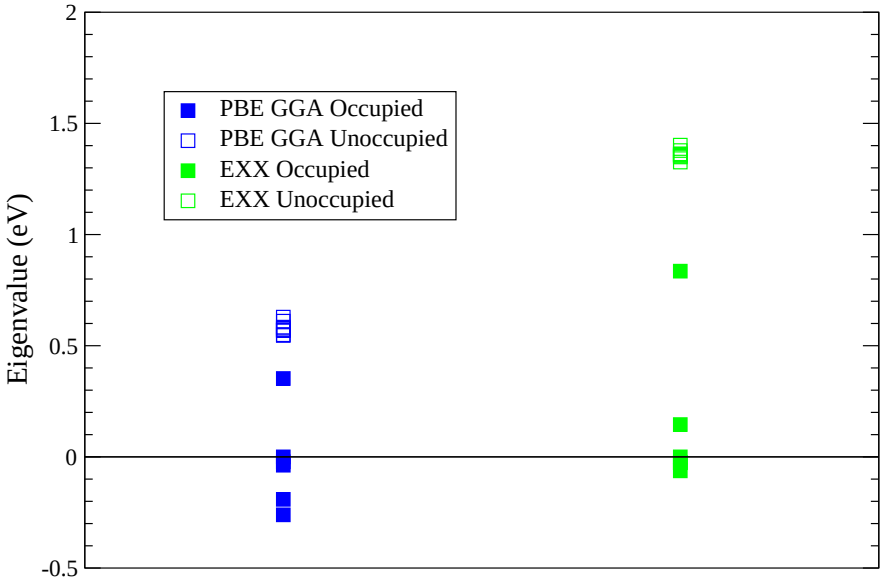


Figure 4.4. The Kohn-Sham eigenvalues obtained from a 217 atom calculation for the -2 charge state of the silicon self-interstitial using EXX and the PBE. The eigenvalues are given relative to the valence band edge.

Chapter 5

Pseudopotentials

One challenge when using a new density functional is finding a way of generating pseudopotentials consistent with that functional. It is possible, of course, to use pseudopotentials developed with another functional, for example LDA. However, such use is not advisable, as the core density replaced by an LDA pseudopotential will not be exactly the same as that replaced by an EXX pseudopotential. We have therefore sought to find a way to generate pseudopotentials consistent with our EXX density functional.

We considered several options for generating EXX-consistent pseudopotentials. First, we considered using one of the wide range of Hartree-Fock pseudopotentials in current use. Such an option is potentially possible, however Hartree-Fock codes do not typically derive the separable pseudopotentials required by the Socorro program, and it was not known whether there were fundamental inconsistencies with the very non-local Hartree-Fock Hamiltonian and the separability algorithms. Secondly, we considered using the FHI98PP software [21] that is in wide use in the DFT community and that contains a KLI functional. The KLI functional is an approximation to the full EXX functional, and might possibly produce core densities arbitrarily close to the EXX core densities. Thirdly, we considered writing our own all-electron DFT code to produce our own core densities and pseudopotentials. We have, in fact, put together an all-electron DFT functionality [22] that we have used for rapid development of algorithms for EXX potential generation. Using this capability for pseudopotential generation remains an option that could be pursued at some point. Finally, we considered obtaining existing software to generate these potentials. We obtained code from the Goerling group [23] and code from the Engel group [24]. After an initial assessment of the two programs, we decided that the code from the Engel Group was more mature, and have used that program to generate EXX pseudopotentials.

Figure 5.1 shows a comparisons for the Si atom of LDA and KLI pseudopotentials constructed with the FHI98PP and the EXX pseudopotential constructed with the Engel group's program. We are currently in the process of testing these potentials in the Socorro program to insure they reproduce meaningful physical properties.

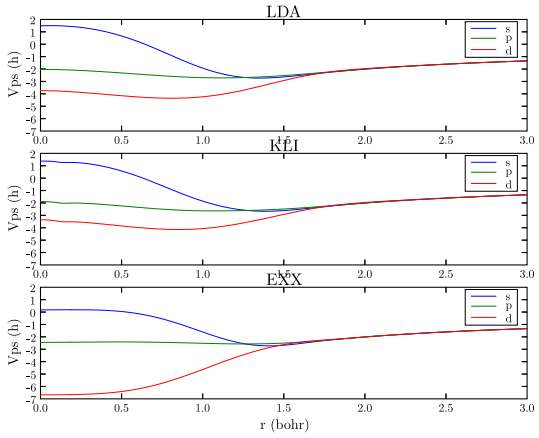


Figure 5.1. A comparison of Si pseudopotentials computed (top) using the Fritz-Haber pseudopotential code with the LDA XC functional, (center) using the Fritz-Haber pseudopotentialcode with the KLI XC functional, (bottom) using the Engels group's pseudopotential code with the EXX XC functional.

Chapter 6

Small system tests of the OEP

We now report the performance of EXX methods on a variety of test systems. We report a comparison of a spectrum of excitation energies for Helium and Beryllium atoms, we report the accuracy of the functionals for molecular hydrogen geometries with stretched H–H bond lengths, and we report the behavior of the band gap of Helium atom at elevated temperatures. These test problems have been computed using an all-electron EXX method described in reference [22].

Helium Atom

Figure 6.1 shows a comparison of Helium excitation energies to the (essentially exact) QMC-derived exchange-correlation functional as well as to HF and to other LDA, GGA, and Hybrid functionals. The two methods derived in the current LDRD project, labelled *EXX* and *EXX-GGA* in the figure, perform significantly better than the HF method and the other density functionals. The average error of the EXX (0.016 h) and the EXX-GGA (0.014 h) methods is significatly lower than that from HF (0.107 h), LDA (0.241 h), GGA (0.227–0.232 h) or hybrid (0.160 h) methods. Furthermore, the EXX values differ from the QMC-derived values by essentially constant values across the entire spectrum, whereas the errors in the other methods fluctuate considerably. As a result, the EXX methods are the only ones that properly predict the ordering of the excited states.

Beryllium Atom

Figure 6.2 reports a similar comparison for Beryllium atom. On the average the HF excitation energies differ by 0.075 h, the LDA, BLYP, and PBE values differ by 0.050–0.070, and the B3LYP values differ by 0.038 h. In contrast, the EXX values differ by only 0.004 h, nearly a factor of ten smaller average difference than the best of the standard DFTs. Moreover, as was seen in He, the OEP excitation energies differ from the QMC values by a constant shift, whereas the LDA, GGA, and HF values fluctuate much more about their averages. The figure illustrates that not only does the EXX perform better on the average, but that each individual excitation value once again differs from the exact levels by a small, nearly constant shift.

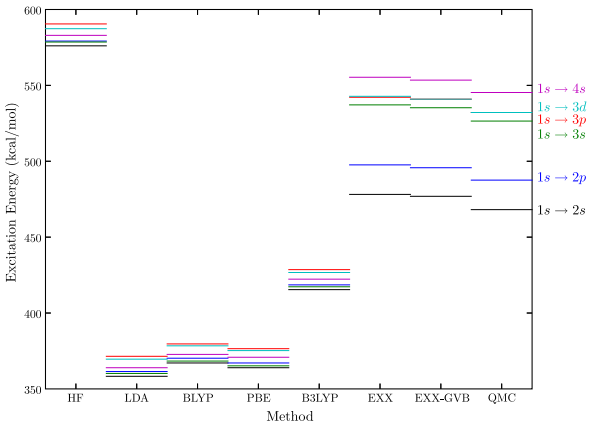


Figure 6.1. A comparison of the excitation energies for different states of Helium atom, using the Hartree-Fock (HF) method, and density functional theory using the LDA, BLYP, PBE, B3LYP functionals, as well as the EXX and EXX-GVB method develop in this LDRD project.

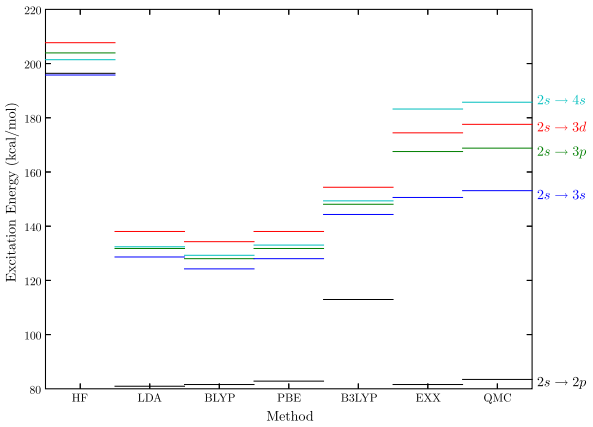


Figure 6.2. A comparison of the excitation energies for different states of Beryllium atom, using the Hartree-Fock (HF) method, and density functional theory using the LDA, BLYP, PBE, B3LYP functionals, as well as the EXX method develop in this LDRD project.

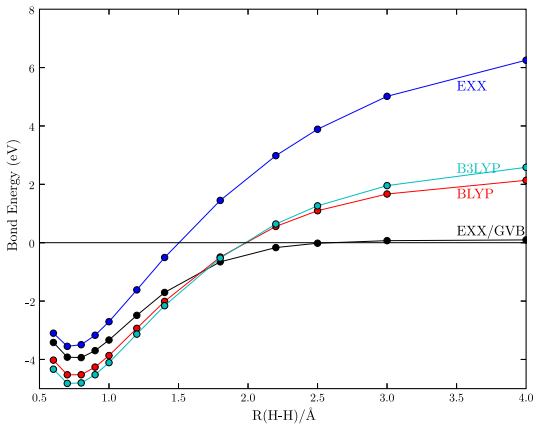


Figure 6.3. A comparison of the dissociation behavior of different density functionals showing that traditional gradient-corrected density functionals (BLYP), hybrid functionals (B3LYP), and even simple OEP/EXX functionals produce energies that are incorrect for molecular hydrogen at geometries with stretched H–H bonds. In contrast, a method from the current work (EXX/GVB) does in fact produces correct energies for the stretched hydrogen geometries.

Hydrogen Molecule

The limitations of the Hartree-Fock wave function for describing partially bonded electronic configurations are well known: the single electronic configuration limits the ways the wave function can be variationally minimized, and thus the energies from Hartree-Fock wave functions do not approach twice the energy of a single H atom as they should. Figure 6.3 shows energies (eV) for molecular hydrogen as the H-H distance (Å) is increased. Several things are notable about this figure. First, the GGA and Hybrid DFT approaches also dissociate to the incorrect limit, even though, being DFT approaches, they don't necessarily suffer from the same limitations. Furthermore, the EXX approach alone dissociates to the same limit that the HF approach would. In the examples above with He and Be atoms, using EXX "fixed" many of the inconsistencies of using a HF wave function, but this is not the case for dissociating chemical bonds. Thirdly, we see that a solution comes from using not a pure Hartree-Fock method in our optimized effective potential optimization, but a correlated wave function approach [22, 25], which then does dissociate to the proper limit.

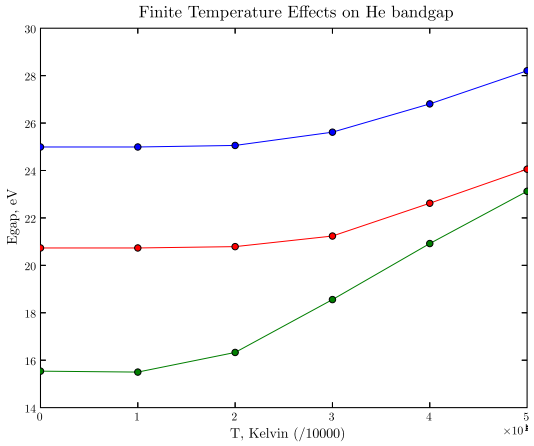


Figure 6.4. A comparison of the band gaps versus temperature using the Hartree-Fock (blue), LDA (green) and EXX (red) methods. Note that the band gap of the EXX method approaches the LDA method at sufficiently high temperatures.

Finite Temperature EXX Simulations

One of our primary motivations for developing an exact exchange capability for DFT calculations as a tool for correcting the systematic errors in the DFT band-gap in finite-temperature molecular dynamics calculations. Thus, realistic performance at finite-temperature is very important to our work. Figure 6.4 shows a comparison of the HF, LDA, and EXX band gaps for Helium atom. Hartree-Fock methods systematically overestimate the band gap, and LDA methods systematically underestimate the band gap. The results from Figure 6.1 suggest that at zero temperture the EXX results are very close to the correct values. Figure 6.4 shows that as we raise the temperature the EXX results approach the LDA results, which are correct at sufficiently high temperature. This study demonstrates that EXX band gaps have the correct behavior at low and at high temperatures.

Chapter 7

Correlation compatible with Exact Exchange

In order to explain the necessity of adding a compatible correlation to exact exchange (EXX) we will use the jellium surface model system in an illustrative example. This model is obtained by performing a self-consistent LDA calculation on a system where electrons move in an external potential of a uniform positively charged background in halfspace. [26] The model itself can be seen as a well-defined procedure to produce densities from which all other quantities can be determined using the Kohn-Sham equations [27] with the LDA exchange-correlation functional. The non-homogeneous densities that define the model are expressed in the dimensionless parameter $r_s = a_0(3/(4\pi)\bar{n})^{1/3}$, where $\bar{n}$ is the density of the positive background, and a_0 is the bohr radius. The total system of positive background and electrons is neutral and the bulk electron density, n_{bulk}, far inside the half space positive background is then equal to $\bar{n}$. Typical jellium surface model densities are shown in Fig. 7.1.

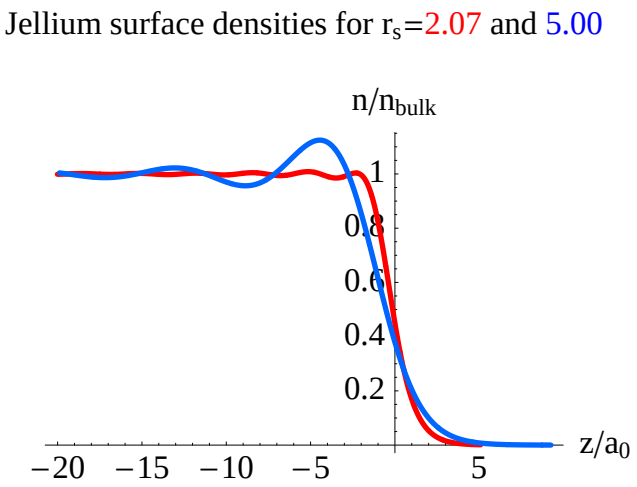


Figure 7.1. Examples of two jellium surface densities with different r_s-values. $r_s = 2.07$ corresponds to the valence electron density in aluminum. Larger r_s corresponds to lower density.

Table 7.1. Exchange surface energies, in erg/cm^2, for the jellium surface model. Mean absolute relative errors (mare) are compared to the EXX results [28].

r_s	LDA	PBE	EXX
2.00	3036	2436	2624
2.07	2673	2125	2296
2.30	1808	1393	1521
2.66	1051	769	854
3.00	668	464	526
3.28	477	315	364
4.00	222	128	157
5.00	92	40	57
mare	**29**%	**13**%	–

For a given jellium surface, that is, density, the only part of the energy that changes using different functionals, $\varepsilon_{xc}(\mathbf{r};[n])$, is the exchange-correlation part, so we only examine this part of the energy. An exchange-correlation surface energy is defined by subtracting the energy an equal amount of electrons would have had in a uniform electron gas with the bulk density (this is equivalent to the LDA expression):

$$\sigma_{xc} = \int n(z)[\varepsilon_{xc}(\mathbf{r};[n]) - \varepsilon_{xc}^{\mathrm{LDA}}(n_{\mathrm{bulk}})]dz \; . \tag{7.1}$$

This expression can be divided into an exchange and a correlation part, $\sigma_{xc} = \sigma_x + \sigma_c$. This division is arbitrary, since only the full σ_{xc} is a physical quantity, not the separate σ_x and σ_c. We will here follow the standard division and define σ_x as the quantity obtained by insering the EXX functional into Eqn. 7.1. The correlation part is defined as 'the rest' and need to be derived with methods outside of DFT. Table 7.1 shows the jellium surface model surface exchange energies for the LDA, PBE, and EXX exchange functionals. The EXX results are from Ref. [28]. EXX being the correct result, the two other functionals are compared to the EXX results by calculation of the mean absolute relative error, mare. As is seen, LDA deviates strongly from the correct answer with 29% mare, while PBE is doing better with 13% mare. Taken out of perspective this discrepancy would be sufficient to motivate the use of the much more computationally expensive EXX functional. However, as pointed out before, the physical quantity is the full exchange-correlation energy, not the separate exchange or correlation energies.

In Tab. 7.2, the surface correlation energies for the same jellium surface model is given. Here the RPA+ is the best available approximation for the exact correlation energy [29]. As seen, also the correlation is far off, in particular for LDA with a 63% mare from the RPA+ results. However, combining exchange and correlation into the real physical quantity, results are given in Tab. 7.3, the errors made in exchange and correlation are cancelling to the extent that LDA only has a 2% error and actually is better than PBE, which had both the separate exchange and correlation better than LDA. This effect is due to the LDA correlation being 'compatible' with LDA exchange. Both the LDA exchange and correlation are derived from a model system, the uniform electron gas, for which LDA thus is the correct answer. Applying LDA to a non-uniform electron gas, such as the

Table 7.2. Correlation surface energies, in erg/cm^2, for the jellium surface model. Mean absolute relative errors (mare) are compared to the RPA+ results [29].

r_s	LDA	PBE	RPA+
2.00	317	827	789
2.07	287	754	719
2.30	210	567	539
2.66	136	381	360
3.00	95	275	255
3.28	72	215	199
4.00	39	124	111
5.00	19	67	56
mare	**63**%	**8**%	–

jellium surface densities in Fig. 7.1, errors made in exchange has a tendency to cancel with errors made in the correlation as is seen in Tabs 7.1-7.3. PBE is not constructed from a model system and thus does not have compatible exchange and correlation, see the tables.

Table 7.3. Exchange-correlation surface energies, in erg/cm^2, for the jellium surface model. Mean absolute relative errors (mare) are compared to the EXX/RPA+ results.

r_s	LDA	PBE	EXX/RPA+
2.00	3354	3264	3413
2.07	2960	2879	3015
2.30	2019	1960	2060
2.66	1187	1150	1214
3.00	763	739	781
3.28	549	530	563
4.00	261	252	268
5.00	111	107	113
mare	**2**%	**5**%	–

So, what correlation should be used for EXX? Let us examine the jellium surface model again. If we use no correlation at all, thus approximating the EXX/RPA+ numbers in Tab. 7.3 with the EXX numbers in Tab. 7.1, we would make errors between 23 and 50%, with a mare of 33%. Combining the EXX in Tab. 7.1 with LDA/PBE correlation in Tab. 7.2 would make errors between 14/1 and 33/10%, mare 21/3%. It would thus be reasonable to use PBE correlation together with EXX. This exercise shows that the usual practice of adding no or LDA correlation to EXX results is not well founded.

We will now proceed and use the jellium surface model system to construct a partially compatible

correlation for EXX by parametrizing the RPA+ correlation in Tab. 7.2. We will follow approximately the same procedure as in [30]. We will use an index dependent on the dimensionless density gradient parameter $s^2 = |\nabla n|^2/(4(3\pi^2)^{2/3}n^{8/3})$ to interpolating between 'interior' ($s = 0$, LDA correlation, index=1) and edge region ($s \to \infty$, index=0):

$$XIDX(s^2) = 1 - cs^2/(cs^2+1) \tag{7.2}$$
$$FIDX(s^2) = 2/(1+exp(s^2/t)) \tag{7.3}$$

These two possible indices are shown in Fig. 7.2, for the values of c and t obtained in the fits described below. The $XIDX$ index is used in the am05 functional[30].

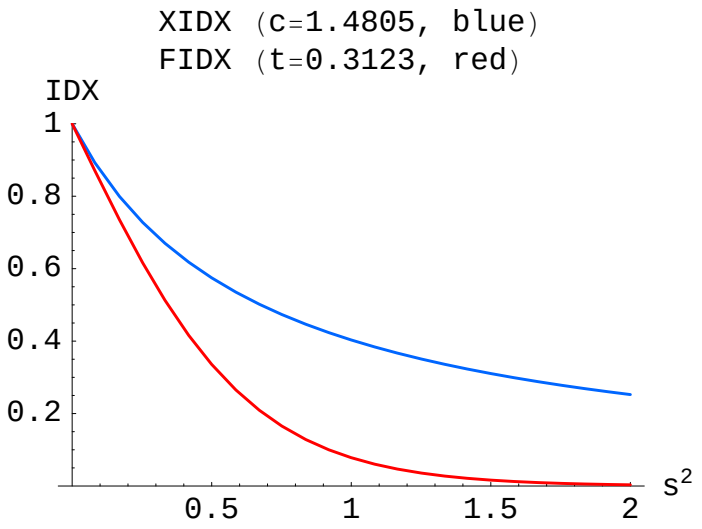


Figure 7.2. Examples of two possible indices to interpolate between 'interior' (LDA correlation, index=1) and edge region (index=0) .

For the interior region, far from the surface, where the density is slowly varying (see Fig. 7.1), LDA correlation (LDAc) is a suitable functional. For the edge region we will use a scaled LDA correlation, with the scaling set by fitting. The two correlation functionals we here construct thus have the following form:

$$EXXc(n,s) = IDX(s^2) * LDAc(n) + (1 - IDX(s^2)) * (g_1 + g_2 * r_s(n)) * LDAc(n)\ , \tag{7.4}$$

where IDX is either $XIDX$ or $FIDX$ from Eqn. 7.3, and g_1 and g_2 are fitting constants. Note also that r_s here is calculated from the local density, not the bulk/positive background density as above. In the construction of am05 [30] only one scaling constant, g_1, was used, but to accurately represent the RPA+ data in Tab. 7.2, two fitting constants are needed. Also note that we correctly would obtain the LDAc for a uniform electron gas system where $s^2 = 0$, since $IDX(0) = 1$ by construction.

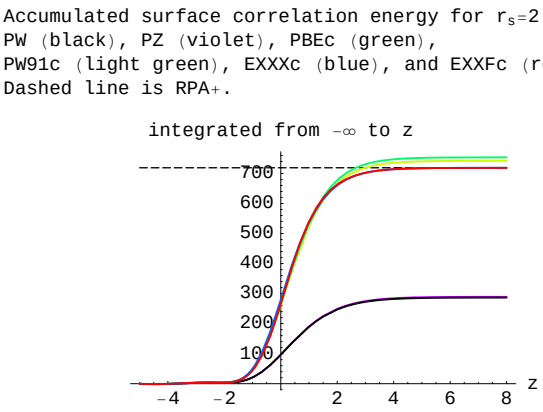
The constants, (g_1, g_2, c) for $EXXXc$ and (g_1, g_2, t) for $EXXFc$, are fitted by non-linear fitting to the RPA+ data in Tab. 7.2. In this case we use the PW parametrization [31] of the LDAc. The PZ correlation [32] would be equally suitable but would give slightly different values of the fitting parameters. Both forms of the index give equally good fits and reproduces the RPA+ results in Tab. 7.2 to within error bars (± 1 erg/cm^2). The constants take the values

$$EXXXc \;:\; (g_1, g_2, c) = (0.02245, 0.03672, 1.4805)\,, \tag{7.5}$$

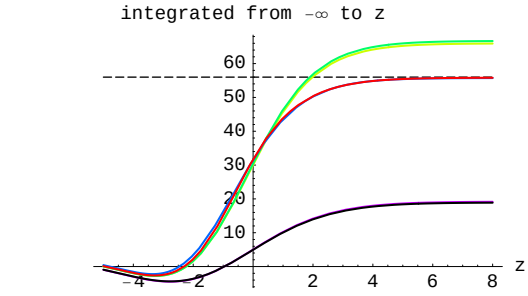
$$EXXFc \;:\; (g_1, g_2, t) = (0.3060, 0.04108, 0.3123)\,. \tag{7.6}$$

In order to compare the two here created correlation functionals with GGA correlations (PBEc [33] and PW91c [34]) and the PW [31] and PZ [32] parametrizations of Ceperley and Alder's Monte Carlo derived LDA correlation [35], we examine the accumulated surface correlation energy obtained from the correlation part of the integrand in Eqn. 7.1 integrated from $-\infty$ to z. This accumulated surface correlation energy is shown in Fig. 7.3 for the same two r_s values as in Fig. 7.1. It is clearly seen that scaling the LDA correlation in the edge region makes $EXXXc$ and $EXXFc$ behave like the GGA correlations even though the derivation of the here presented correlation functionals has nothing in common with the derivation of the PBEc and PW91c. We also see that the main difference from the GGA correlations is picked up outside of the surface. It is also interesting to note that despite that the indices in Eqn. 7.3 are quite different (see also Fig. 7.2) the final results are very similar. A final, obvious, observation is that the accumulated surface correlation indeed converge towards the RPA+ values that the full integrals were fitted to.

We have constructed two correlation functionals that reproduce the correct correlation in two model systems, the uniform electron gas and the jellium surface system. This correlation is thus partially compatible with EXX. The two correlation functionals seems to reproduce GGA results for moderate values of s^2. Full and realistic testing is needed before any final conclusions about the performance of these two correlation functionals can be made.



(a) $r_s = 2.07$.



(b) $r_s = 5.00$.

Figure 7.3. Accumulated surface correlation energy.

Appendix A

Expressions for the Exchange Energy and Exchange Derivative

Implementation of the exact-exchange, optimized-effective-potential technique in Socorro required us to write subroutines for computing the exchange energy and the derivative of the exchange energy with respect to a Kohn-Sham function. A significant challenge was to deal with the integrable divergences that arise in the expressions for these quantities. The expression for the exchange energy is derived first below and we discuss the treatment of the integrable divergence. The expression for the derivative of the exchange energy is then derived further below. In these derivations, $\rho(\mathbf{r},\mathbf{r}')$ is the one electron density matrix at $\mathbf{r}$ and $\mathbf{r}'$,

$$\rho_1(\mathbf{r},\mathbf{r}') = e\sum_{\substack{n\\k}} f_{n,\mathbf{k}}\phi^*_{n,\mathbf{k}}(\mathbf{r})\phi_{n,\mathbf{k}}(\mathbf{r}'),$$

where $\phi_{n,\mathbf{k}}(\mathbf{r})$ is a Kohn-Sham function,

$$\phi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$$

$f_{n,\mathbf{k}}$ is its occupation, and e is the electron charge which is equal to $\sqrt{2}$ in Rydberg atomic units. The exchange energy, E_x, is then

$$E_x = -\frac{1}{2}\int \frac{\rho_1(\mathbf{r},\mathbf{r}')\rho_1^*(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3r d^3r'$$

$$= -\frac{e^2}{2}\sum_{\substack{n\\\mathbf{k}}}\sum_{\substack{n'\\\mathbf{k}'}} f_{n,\mathbf{k}} f_{n',\mathbf{k}'}\int \frac{\phi^*_{n,\mathbf{k}}(\mathbf{r})\phi_{n,\mathbf{k}}(\mathbf{r}')\phi_{n',\mathbf{k}'}(\mathbf{r})\phi^*_{n',\mathbf{k}'}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3r d^3r'$$

$$= -\frac{e^2}{2}\sum_{\substack{n\\\mathbf{k}}}\sum_{\substack{n'\\\mathbf{k}'}} f_{n,\mathbf{k}} f_{n',\mathbf{k}'}\int u_{n',\mathbf{k}'}(\mathbf{r})u^*_{n,\mathbf{k}}(\mathbf{r})e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}\left[\int \frac{u^*_{n',\mathbf{k}'}(\mathbf{r}')u_{n,\mathbf{k}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}'}d^3r'\right]d^3r$$

$$= -\frac{e^2}{2}\sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} \left[\int \frac{h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} d^3r' \right] d^3r$$

$$= -\frac{e^2}{2}\sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} \left[\sum_{\boldsymbol{G}} h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) \int \frac{e^{i(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})\cdot\boldsymbol{r}'}}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3r' \right] d^3r$$

$$= -2\pi e^2 \sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} \left[\sum_{\boldsymbol{G}} \frac{h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})}{|\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}|^2} e^{i(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})\cdot\boldsymbol{r}} \right] d^3r$$

$$= -2\pi e^2 \sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} \frac{h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})}{|\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}|^2} e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$= -2\pi e^2 \sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$= -2\pi e^2 \sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) d^3r$$

We next discuss the treatment of the integrable divergence that appears in this expression. In DFT plane wave calculations, $u_{n,\boldsymbol{k}}(\boldsymbol{r})$ is normalized over a cell volume (Ω) and is constrained to be orthogonal to $u_{n',\boldsymbol{k}}(\boldsymbol{r})$ when $n' \neq n$. As such, the integrals of $h_{n'=n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})$ and $h_{n'\neq n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})$ over the cell volume – $h_{n'=n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$ and $h_{n'\neq n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$ – have the values 1 and 0, respectively. In the latter case, the terms $H_{n'\neq n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$ will also be equal to 0 and the numerical problem of dividing by $|\boldsymbol{G}|^2 = 0$ is avoided by skipping them. In the former case, the terms $H_{n'=n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$ are undefined and thus require special treatment. To treat these terms, I begin by separating them from the others:

$$\sum_{\substack{n \\ \boldsymbol{k}}}\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$= \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{\substack{n' \\ \boldsymbol{k}' \neq \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^{*}_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{n'} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}} \int h^{*}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$= \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{\substack{n' \\ \boldsymbol{k}' \neq \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^{*}_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{n' \neq n} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}} \int h^{*}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}} \int h^{*}_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$= \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{\substack{n' \\ \boldsymbol{k}' \neq \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^{*}_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{n' \neq n} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}} \int h^{*}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G} \neq \boldsymbol{0}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}} \int h^{*}_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G} \neq \boldsymbol{0}} H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}} \left[H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G} = 0) \right] \int h^{*}_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) d^3r$$

$$= \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{\substack{n' \\ \boldsymbol{k}' \neq \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} \int h^{*}_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{n'} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}} \int h^{*}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G} \neq \boldsymbol{0}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] d^3r$$

$$+ \sum_{\substack{n \\ \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}} \left[H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G} = 0) \right] \int h^{*}_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r}) d^3r$$

The divergent terms are in the final sum in the expression above. Gygi and Baldereschi [PRB 34, 4405 (1986)] noted similar divergent terms in their expressions for the matrix elements of the Hartree-Fock exchange operator in the reciprocal space

representation. They noted that the divergence is integrable in that the *analytic* integral of $H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$ over the Brillouin zone is finite. Given this, they replaced the divergent terms as shown below:

$$H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0}) = h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left\{\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{k}} F(\boldsymbol{k}-\boldsymbol{k}'')\right\}$$

$$= h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left\{\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right\}$$

where F is a function which is periodic in reciprocal space, has the same type (i.e. $|\boldsymbol{q}|^{-2}$) of divergence as $H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$, and is sufficiently smooth that it can be numerically integrated using a modest number of points in the Brillouin zone. I note that the $\boldsymbol{k}$-point mesh used to perform the (partial) numerical integration of F does not have to be the same mesh on which the wave functions are computed. As such, this replacement allows for calculations at a single $\boldsymbol{k}$-point in the Brillouin zone. The expression given above for the exchange energy is then written as

$$E_x = -2\pi e^2\left\{\sum_{\boldsymbol{n}}\sum_{\substack{\boldsymbol{n}'\\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'}\int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r})\left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\right]d^3r\right.$$

$$+\sum_{\boldsymbol{n}}\sum_{\boldsymbol{n}'} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}}\int h^*_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})\left[\sum_{\boldsymbol{G}\neq\boldsymbol{0}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G})e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\right]d^3r$$

$$\left.+\sum_{\boldsymbol{n}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}} h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left[\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right]\int h^*_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})d^3r\right\}$$

$$= -2\pi e^2\left\{\sum_{\boldsymbol{n}}\sum_{\substack{\boldsymbol{n}'\\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'}\int h^*_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r})H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r})d^3r\right.$$

$$+\sum_{\boldsymbol{n}}\sum_{\boldsymbol{n}'} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}}\int h^*_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})\tilde{H}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})d^3r$$

$$\left.+\sum_{\boldsymbol{n}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}}\left|h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\right|^2\left[\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right]\right\}$$

$$= -2\pi e^2 \left\{ \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{\substack{n' \\ \boldsymbol{k}' \neq \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}'} S_{n',\boldsymbol{k}',n,\boldsymbol{k}} + \sum_{\substack{n \\ \boldsymbol{k}}} \sum_{n'} f_{n,\boldsymbol{k}} f_{n',\boldsymbol{k}} \tilde{S}_{n',\boldsymbol{k},n,\boldsymbol{k}} \right.$$

$$\left. + \sum_{\substack{n \\ \boldsymbol{k}}} f_{n,\boldsymbol{k}} f_{n,\boldsymbol{k}} \left| h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0}) \right|^2 \left[\int_{BZ} F(\boldsymbol{q}) d\boldsymbol{q} - \sum_{\boldsymbol{k}'' \neq \boldsymbol{0}} F(\boldsymbol{k}'') \right] \right\}$$

We next derive the expression for the derivative of the exchange energy with respect to a Kohn-Sham function.

$$\frac{1}{\Omega} \frac{\delta E_x}{\delta \phi^*_{n,\boldsymbol{k}}(\boldsymbol{r})} = -\frac{e}{2} f_{n,\boldsymbol{k}} \int \frac{\rho_1^*(\boldsymbol{r},\boldsymbol{r}') \phi_{n,\boldsymbol{k}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3 r'$$

$$= -\frac{e^2}{2} f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} \phi_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\int \frac{\phi^*_{n',\boldsymbol{k}'}(\boldsymbol{r}') \phi_{n,\boldsymbol{k}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3 r' \right]$$

$$= -\frac{e^2}{2} f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) e^{i\boldsymbol{k}'\cdot\boldsymbol{r}} \left[\int \frac{u^*_{n',\boldsymbol{k}'}(\boldsymbol{r}') u_{n,\boldsymbol{k}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} d^3 r' \right]$$

$$= -\frac{e^2}{2} f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) e^{i\boldsymbol{k}'\cdot\boldsymbol{r}} \left[\int \frac{h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} d^3 r' \right]$$

$$= -\frac{e^2}{2} f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) e^{i\boldsymbol{k}'\cdot\boldsymbol{r}} \left[\sum_{\boldsymbol{G}} h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) \int \frac{e^{i(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})\cdot\boldsymbol{r}'}}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3 r' \right]$$

$$= -2\pi e^2 f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) e^{i\boldsymbol{k}'\cdot\boldsymbol{r}} \left[\sum_{\boldsymbol{G}} \frac{h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})}{|\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}|^2} e^{i(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})\cdot\boldsymbol{r}} \right]$$

$$= -2\pi e^2 f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} \frac{h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})}{|\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}|^2} e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$$

$$= -2\pi e^2 f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$$

$$= -2\pi e^2 f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$$

Then

$$\frac{1}{\Omega}\frac{\delta E_x}{\delta u^*_{n,\boldsymbol{k}}(\boldsymbol{r})} = -2\pi e^2 f_{n,\boldsymbol{k}} \sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r})$$

The treatment of the integrable divergence in this expression proceeds in the same way as noted above for the exchange energy. I begin by separating the divergent term from the others:

$$\sum_{\substack{n' \\ \boldsymbol{k}'}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right]$$

$$= \sum_{\substack{n' \\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] + \sum_{n'} f_{n',\boldsymbol{k}} u_{n',\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right]$$

$$= \sum_{\substack{n' \\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] + \sum_{n'\neq n} f_{n',\boldsymbol{k}} u_{n',\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right]$$
$$+ f_{n,\boldsymbol{k}} u_{n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right]$$

$$= \sum_{\substack{n' \\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] + \sum_{n'\neq n} f_{n',\boldsymbol{k}} u_{n',\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}\neq\boldsymbol{0}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right]$$
$$+ f_{n,\boldsymbol{k}} u_{n,\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}\neq\boldsymbol{0}} H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] + f_{n,\boldsymbol{k}} u_{n,\boldsymbol{k}}(\boldsymbol{r}) H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$$

$$= \sum_{\substack{n' \\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'} u_{n',\boldsymbol{k}'}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right] + \sum_{n'} f_{n',\boldsymbol{k}} u_{n',\boldsymbol{k}}(\boldsymbol{r}) \left[\sum_{\boldsymbol{G}\neq\boldsymbol{0}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \right]$$
$$+ f_{n,\boldsymbol{k}} u_{n,\boldsymbol{k}}(\boldsymbol{r}) H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$$

The divergent term is in the final sum in the expression above. As explained above, Gygi and Baldereschi [PRB 34, 4405 (1986)] noted similar divergent terms in their expressions for the matrix elements of the Hartree-Fock exchange operator in the reciprocal space representation. They further noted that the divergence is integrable in

that the *analytic* integral of $H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$ over the Brillouin zone is finite. Given this, they replaced the divergent term as shown below:

$$H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0}) = h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left\{\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{k}} F(\boldsymbol{k}-\boldsymbol{k}'')\right\}$$

$$= h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left\{\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right\}$$

where F is a function which is periodic in reciprocal space, has the same type (i.e. $|\boldsymbol{q}|^{-2}$) of divergence as $H_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})$, and is sufficiently smooth that it can be numerically integrated using a modest number of points in the Brillouin zone. I note that the $\boldsymbol{k}$-point mesh used to perform the numerical integration of F does not have to be the same mesh on which the wave functions are computed. As such, this replacement allows for calculations at a single $\boldsymbol{k}$-point in the Brillouin zone. The expression given above for the exchange derivative is then written as

$$\frac{1}{\Omega}\frac{\delta E_x}{\delta u^*_{n,\boldsymbol{k}}(\boldsymbol{r})} = -2\pi e^2 f_{n,\boldsymbol{k}}\left\{\sum_{\substack{n'\\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'}u_{n',\boldsymbol{k}'}(\boldsymbol{r})\left[\sum_{\boldsymbol{G}} H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\right]\right.$$
$$+\sum_{n'} f_{n',\boldsymbol{k}}u_{n',\boldsymbol{k}}(\boldsymbol{r})\left[\sum_{\boldsymbol{G}\neq\boldsymbol{0}} H_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G})e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\right]$$
$$\left.+ f_{n,\boldsymbol{k}}u_{n,\boldsymbol{k}}(\boldsymbol{r})h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left[\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right]\right\}$$

$$= -2\pi e^2 f_{n,\boldsymbol{k}}\left\{\sum_{\substack{n'\\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'}u_{n',\boldsymbol{k}'}(\boldsymbol{r})H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) + \sum_{n'} f_{n',\boldsymbol{k}}u_{n',\boldsymbol{k}}(\boldsymbol{r})\tilde{H}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})\right.$$
$$\left.+ f_{n,\boldsymbol{k}}u_{n,\boldsymbol{k}}(\boldsymbol{r})h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left[\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right]\right\}$$

$$= -2\pi e^2 f_{n,\boldsymbol{k}}\left\{\sum_{\substack{n'\\ \boldsymbol{k}'\neq\boldsymbol{k}}} f_{n',\boldsymbol{k}'}D_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r}) + \sum_{n'} f_{n',\boldsymbol{k}}\tilde{D}_{n',\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{r})\right.$$
$$\left.+ f_{n,\boldsymbol{k}}u_{n,\boldsymbol{k}}(\boldsymbol{r})h_{n,\boldsymbol{k},n,\boldsymbol{k}}(\boldsymbol{G}=\boldsymbol{0})\left[\int_{BZ} F(\boldsymbol{q})d\boldsymbol{q} - \sum_{\boldsymbol{k}''\neq\boldsymbol{0}} F(\boldsymbol{k}'')\right]\right\}$$

To complete this section, we comment on the possibility of alaising errors. Note that whereas $u_{n,\boldsymbol{k}}(\boldsymbol{G})=0$ for $|\boldsymbol{G}|>G_{\max}$, $h_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})$ and $H_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})$ can be non-zero for $|\boldsymbol{G}|\le 2G_{\max}$ and $D_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})$ can be non-zero for $|\boldsymbol{G}|\le 3G_{\max}$. Furthermore, the mesh typically used in a DFT plane wave calculation is only able to resolve functions having non-zero Fourier coefficients for $|\boldsymbol{G}|\le 2G_{\max}$. As such, $D_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{r})$ will not be resolved and $D_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})$ will therefore contain aliasing errors for $G_{\max}<|\boldsymbol{G}|\le 2G_{\max}$. This is not a problem in practice because only $D_{n',\boldsymbol{k}',n,\boldsymbol{k}}(\boldsymbol{G})$ for $|\boldsymbol{G}|\le G_{\max}$ are needed to update $u_{n,\boldsymbol{k}}(\boldsymbol{G})$ and these coefficients do not contain aliasing errors. (The same comments apply to LDA and GGA exchange derivatives.)

References

[1] M. P. Desjarlais, J. D. Kress, and L. A. Collins, Phys. Rev. E **66**, 025401 (2002).

[2] M. P. Desjarlais, Phys. Rev. B **68**, 064204 (2003).

[3] S. Mazevet, M. P. Desjarlais, L. A. Collins, J. D. Kress, and N. H. Magee, Phys. Rev. E **71**, 016409 (2005).

[4] T. R. Mattsson and M. P. Desjarlais, Phys. Rev. Lett. **97**, 017801 (2006).

[5] W. L. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[6] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[7] J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).

[8] V. Sahni, J. Gruenebaum, and J. P. Perdew, Phys. Rev. B **26**, 4371 (1982).

[9] E. Engel and S. H. Vosko, Phys. Rev. A **47**, 2800 (1993); Phys. Rev. B **50**, 10498 (1994).

[10] A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994).

[11] T. Kotani, Phys. Rev. Lett. **74**, 2989 (1995).

[12] M. Städele, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. Lett. **79**, 2089 (1997); M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10031 (1999).

[13] A. Görling, Phys. Rev. Lett. **83**, 5459 (1999).

[14] R. Hyman, M. Stiles, , and A. Zangwill. A gradient search method for orbital-dependent charge-density and current density functional calculations. *Phys. Rev. B*, 62(23):15521–15526, 2000.

[15] S. Kümmel and J. P. Perdew. Simple iterative construction of the optimized effective potential for orbital functionals, including exact exchange. *Phys. Rev. Letters*, 90(4):043004, 2003.

[16] R. A. Lippert, N. A. Modine, and A. F. Wright. The optimized effective potential with finite temperature. *J. Phys.: Condens. Matter*, 18(17):4295, 2006.

[17] R. A. Horn and C. R. Johnson. *Topics in Matrix analysis*. Cambridge University Press, Cambridge, 1991.

[18] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[19] G. H. Golub and C. F. V. Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 2d edition, 1989.

[20] V. N. Staroverov, G. E. Scuseria, and E. R. Davidson. Optimized effective potentials yielding Hartree-Fock energies and densities. *J. Chem. Phys.*, 124(14):141103, 2006.

[21] Martin Fuchs and Matthias Scheffler. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density functional theory. *Comp. Phys. Comm.*, 119:67–98, 1999.

[22] Richard P. Muller and Michael P. Desjarlais. Optimized effective potential from a correlated wave function: Optimized effective potential-generalized valence bond (OEP-GVB). *J. Chem. Phys.*, 125:54101, 2006.

[23] M. Moukara, M. Stadele, J. A. Majewski, P. Vogl, and A. Goerling. Norm-conserving pseudopotentials in the exact-exchange Kohn-Sham formalism. *J. Phys.: Condens. Matter*, 12:6783–6798, 2000.

[24] E. Engel, A. Hock, R. N. Schmid, R. M. Dreizler, and N. Chetty. Role of the core-valence interaction for pseudopotential calculations with exact exchange. *Phys. Rev. B*, 64:125111, 2001.

[25] Frank W. Bobrowicz and William A. Goddard, III. The self-consistent field equations for generalized valence bond and open-shell Hartree-Fock wave functions. In Henry F. Schaefer, III, editor, *Methods of Electronic Structure Theory*, volume 3 of *Modern Theoretical Chemistry*, page 79. Plenum Press, New York, 1977.

[26] N. D. Lang and W. Kohn, "Theory of Metal Surfaces: Charge Density and Surface Energy", Physical Review B **1**, 4555 (1970).

[27] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Physical Review **140**, A1133 (1964).

[28] J. M. Pitarke and A. G. Eguiluz, "Jellium surface energy beyond the local-density approximation: Self-consistent-field calculations", Physical Review B **63**, 045116 (2001).

[29] Z. Yan, J. P. Perdew, and S. Kurth, "Density functional for short-range correlation: Accuracy of the random-phase approximation for isoelectronic energy changes", Physical Review B **61**, 16430 (2000).

[30] R. Armiento and A. E. Mattsson, "Functional designed to include surface effects in self-consistent density functional theory", Physical Review B **72**, 085108 (2005).

[31] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", Physical Review B **45**, 13244 (1992).

[32] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", Physical Review B **23**, 5048 (1981).

[33] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", Physical Review Letters **77** 3865 (1996).

[34] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singhi, and C. Fiolhais, "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation", Physical Review B **46**, 6671 (1992).

[35] D. M. Ceperley and B. J. Alder, "Ground State of the Electron Gas by a Stochastic Method", Physical Review Letters **45**, 566 (1980).

DISTRIBUTION:

1	MS 1181	Tom Mehlhorn, 1640
2	MS 1186	Mike Desjarlais, 1640
1	MS 1322	Rick Muller, 1435
1	MS 1322	Ann Mattsson, 1435
1	MS 1322	Mark Sears, 1435
1	MS 1415	Alan Wright, 1112
1	MS 1415	Normand Modine, 1132
2	MS 9018	Central Technical Files, 8944
2	MS 0899	Technical Library, 4536
1	MS 0123	D. Chavez, LDRD Office, 1011

Sandia National Laboratories