The optimized effective potential with finite temperature

R A Lippert[1], N A Modine[2] and A F Wright[2]

[1] MIT Department of Mathematics, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA
[2] Sandia National Laboratories, PO Box 5800, Albuquerque, NM 87185-0601, USA

E-mail: lippert@math.mit.edu, namodin@sandia.gov and afwrigh@sandia.gov

Abstract
The optimized effective potential (OEP) method provides an additional level of
exactness in the computation of electronic structures, e.g. the exact exchange
energy can be used. This extra freedom is likely to be important in moving
density functional methods beyond traditional approximations such as the local
density approximation. We provide a new density-matrix-based derivation of
the gradient of the Kohn–Sham energy with respect to the effective potential.
This gradient can be used to iteratively minimize the energy in order to find the
OEP. Previous work has indicated how this can be done in the zero temperature
limit. This paper generalizes the previous results to the finite temperature
regime. Equating our gradient to zero gives a finite temperature version of the
OEP equation.

1. Introduction

The Kohn–Sham density functional theory (DFT) [1] 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 the local density approximation (LDA) and the generalized gradient
approximation (GGA), take E_{XC} to be an explicit functional of the electronic density. Since
the non-interacting Kohn–Sham orbitals are implicit functionals of the electronic density [2],
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 [3–9]. An explicit dependence on the orbitals allows

approximate E_{XC} expressions to capture physical behaviours of the exact Kohn–Sham E_{XC} that cannot 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 behaviour of the exact exchange energy (see, for example, [10]). The exact exchange energy is the essential element in achieving the fourth rung of Perdew's famous Jacob's ladder of density functional approximations [11].

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 cannot 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 [3–5, 7], or by directly evaluating and inverting a response function [6, 8, 9].

Several recent papers have proposed calculation of the OEP by means of an iterative minimization of the energy [12–15]. Yang and Wu expanded the potential as a sum of basis functions, and used the chain rule and perturbation theory to derive an expression for the derivative of the energy with respect to the coefficients of the basis functions [12]. Their expression involves a sum over unoccupied orbitals, and therefore their approach is most suitable for basis sets, such as Gaussian orbitals, that use a relatively small number of functions. Hyman, Stiles and Zangwill used Lagrange multiplier methods to derive an expression for the gradient of the energy with respect to the potential that does not involve an explicit sum over unoccupied orbitals, and they proposed using this gradient to minimize the energy iteratively [13]. Kummel and Perdew derived an expression nearly identical to that of Hyman, Stiles and Zangwill [14, 15]. 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 the expression found in [13–15] 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 [16, 17]. 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 ability to perform calculations with advanced functionals that have an explicit dependence on the orbitals at non-zero temperature is particularly important 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 paper 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 DFT, we apply the results of section 2 to the density matrix P viewed as a function of the Kohn–Sham Hamiltonian H, and thereby derive a finite temperature OEP equation in terms of H and P. 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.

2. Perturbation theory of operator functions

A scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ naturally extends to a mapping of Hermitian operators to Hermitian operators by

$$f(H) = \sum_i f(\lambda_i)\phi_i\phi_i^\dagger, \tag{1}$$

summing over all eigenvectors ϕ_i of H with eigenvalue λ_i (see [18], chapter 6). The example with which we are primarily concerned is when H is a finite sized Hamiltonian matrix and $f(x) = (1 + \exp(\beta x))^{-1}$ is the Fermi function, although the results in this section are more general. Our primary goal is to develop formulas for the first variation of $f(H)$ with respect to an arbitrary variation in H.

If $\phi_1, \ldots, \phi_{N_b}$ is a complete basis of eigenvectors of H, with eigenvalue λ_i, and f is analytic, then the first variation (by an infinitesimal δH) has a simple form,

$$\delta f(H) = f(H + \delta H) - f(H) = \sum_{i,j=1}^{N_b} \Delta f(\lambda_i, \lambda_j)\phi_i \left(\phi_i^\dagger \delta H \phi_j \right) \phi_j^\dagger$$

where $\Delta f(x, y)$ is the first divided difference of f, given by

$$\Delta f(x, y) = \begin{cases} \dfrac{f(x) - f(y)}{x - y} & x \neq y \\ f'(x) & x = y. \end{cases}$$

This formula can be directly verified for the matrix-power functions H^k for all k,

$$(H + \delta H)^k - H^k = \sum_{m=0}^{k-1} H^m(\delta H)H^{k-m-1}$$

$$= \sum_{m=0}^{k-1} \sum_{ij} \lambda_i^m \lambda_j^{k-m-1} \phi_i \left(\phi_i^\dagger \delta H \phi_j \right) \phi_j^\dagger$$

$$= \sum_{ij} \begin{cases} \dfrac{\lambda_i^k - \lambda_j^k}{\lambda_i - \lambda_j} & \lambda_i \neq \lambda_j \\ k\lambda_i^{k-1} & \lambda_i = \lambda_j \end{cases} \phi_i \left(\phi_i^\dagger \delta H \phi_j \right) \phi_j^\dagger,$$

which, by linearity, extends to all power series.

3. Density functional theory review

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

$$E(P) = \text{tr}\{P(K + V_I)\} + E_{\text{HXC}}(P) - \frac{1}{\beta} S(P). \tag{2}$$

The unconstrained derivative is

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

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, P is the minimizer of $\mathrm{tr}\{PH\} - (1/\beta)S(P)$ with $\mathrm{tr}\{P\} = n$, which is equivalent to the conditions

$$P = \frac{1}{1 + \exp(\beta(H - \mu I))} = f_\beta(H - \mu I) \tag{3}$$

$$\mathrm{tr}\{P\} = n \tag{4}$$

for some chemical potential μ. Thus, we can consider P to be parametrized by two unknowns V and μ with two relations (3) and (4). Note: one could absorb μ into V, but we find it advantageous to keep it distinct in its role as a Lagrange multiplier.

With P satisfying these relations, the energy differential simplifies to

$$\frac{\partial E}{\partial P} = K + V_I + \frac{\partial E_{\mathrm{HXC}}}{\partial P} + \frac{1}{\beta} \log(P(I - P)^{-1}) \tag{5}$$

$$= K + V_I + \frac{\partial E_{\mathrm{HXC}}}{\partial P} - (H - \mu I) \tag{6}$$

$$= \frac{\partial E_{\mathrm{HXC}}}{\partial P} - (V - \mu I). \tag{7}$$

4. Finite temperature OEP with density operators

From section 3, the density matrix is related to the Kohn–Sham Hamiltonian, H, by (3) and (4). Let ε_i be the eigenvalues of H, and let $\omega_i = f_\beta(\varepsilon_i - \mu)$ be the eigenvalues of P. The divided difference factors are given by

$$\omega_{ij} \equiv \Delta f_\beta(\varepsilon_i - \mu, \varepsilon_j - \mu) = \begin{cases} \dfrac{\omega_i - \omega_j}{\varepsilon_i - \varepsilon_j} & \varepsilon_i \neq \varepsilon_j \\ -\beta\omega_i(1 - \omega_i) & \varepsilon_i = \varepsilon_j \end{cases} \tag{8}$$

and we write $\Delta\omega[X] = \sum_{ij} \omega_{ij}\phi_i(\phi_i^\dagger X \phi_j)\phi_j^\dagger$. In an H-diagonalizing basis, with $X_{ij} \equiv \phi_i^\dagger X \phi_j$ for any X, $\Delta\omega$ has the action $X_{ij} \to \omega_{ij}X_{ij}$. We note that

$$\omega_{ij} = \begin{cases} -\dfrac{\sinh(\beta(\varepsilon_i - \varepsilon_j)/2)}{2(\varepsilon_i - \varepsilon_j)\cosh(\beta(\varepsilon_i - \mu)/2)\cosh(\beta(\varepsilon_j - \mu)/2)} & \varepsilon_i \neq \varepsilon_j \\ -\dfrac{\beta}{4\cosh(\beta(\varepsilon_i - \mu)/2)^2} & \varepsilon_i = \varepsilon_j \end{cases}$$

is equivalent to (8), but numerically more accurate.

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

$$\delta E(H) = \mathrm{tr}\left\{\frac{\partial E(P(H, \mu))}{\partial P}\delta P\right\}$$

in an H-diagonalizing basis,

$$\delta P_{ij} = \omega_{ij}\left(\delta H_{ij} - \delta\mu\delta_{ij}\right).$$

By (4), the trace of δP vanishes,

$$\delta\mu = \frac{\sum_{ii} \omega_{ii}\delta H_{ii}}{\sum_{ii} \omega_{ii}} = \frac{\mathrm{tr}\{P(I - P)\delta H\}}{\mathrm{tr}\{P(I - P)\}}.$$

Thus, in an H-diagonalizing basis,

$$
\begin{aligned}
\delta E &= \sum_{ij} \frac{\partial E}{\partial P_{ij}} \omega_{ij} \left(\delta H_{ij} - \delta \mu \delta_{ij} \right) \\
&= \sum_{ij} \frac{\partial E}{\partial P_{ij}} \omega_{ij} \left(\delta H_{ij} - \delta_{ij} \frac{\sum_k \omega_{kk} \delta H_{kk}}{\sum_k \omega_{kk}} \right) \\
&= \sum_{ij} \omega_{ij} \left(\frac{\partial E}{\partial P_{ij}} - \delta_{ij} \frac{1}{\sum_k \omega_{kk}} \sum_k \omega_{kk} \frac{\partial E}{\partial P_{kk}} \right) \delta H_{ij}
\end{aligned}
$$

and the gradient is therefore

$$
\frac{\partial E}{\partial H_{ij}} = \omega_{ij} \left(\frac{\partial E}{\partial P_{ij}} - \delta_{ij} \frac{\chi}{\zeta} \right) \tag{9}
$$

$$
\frac{\partial E}{\partial H} = \Delta \omega \left[\frac{\partial E}{\partial P} - \frac{\chi}{\zeta} I \right] \tag{10}
$$

where $\chi = \sum_{kk} \omega_{kk} \frac{\partial E}{\partial P_{kk}} = \text{tr}\left\{ \Delta \omega \left[\frac{\partial E}{\partial P} \right] \right\}$ and $\zeta = \sum_k \omega_{kk} = \text{tr}\{\Delta \omega [I]\}$. Note: these expressions are manifestly traceless.

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

$$
\frac{\partial E}{\partial V(r)} = \sum_{ij} \phi_i(r) \frac{\partial E}{\partial H_{ij}} \phi_j^*(r) \tag{11}
$$

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

5. Finite temperature OEP with orbitals

Instead of representing the density matrix, it is often more practical to express P in terms of an incomplete basis of partially occupied orbitals. Let $\phi_1, \phi_2, \ldots, \phi_{N_b}$ be a complete eigenbasis of H sorted non-decreasingly in eigenvalue (H is $N_b \times N_b$). Let N be sufficiently large that $\omega_i, \omega_{ij} \sim 0$ when $i, j > N$. Then we may truncate the basis so that

$$
P = \phi \Omega \phi^\dagger \tag{12}
$$

where $\phi = [\,\phi_1 \quad \phi_2 \quad \cdots \quad \phi_N\,]$ is $N_b \times N$, Λ is the diagonal matrix of the first N eigenvalues and Ω is the diagonal matrix with entries $\omega_1, \ldots, \omega_N$ (i.e. $\Omega = f_\beta(\Lambda - \mu I)$). Note: N will usually be smaller than the number of primitive basis elements, N_b, 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 $N_b \times N_b$ orthogonal projector onto the span of the ϕ_i (and hence commutes with H).

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

$$
\chi = \sum_{i=1}^N \omega_{ii} \frac{\partial E}{\partial P_{ii}}, \qquad \zeta = \sum_{i=1}^N \omega_{ii}.
$$

Since $\omega_{ij} \sim 0$ when $i, j > N$, all components of $\frac{\partial E}{\partial H_{ij}}$, given by (9), likewise vanish and

$$
\begin{aligned}
\frac{\partial E}{\partial H} &= \sum_{i,j \leqslant N} \phi_i \bar{J}_{ij} \phi_j^\dagger + \sum_{i \leqslant N < j} \phi_i \left(\omega_{ij} \frac{\partial E}{\partial P_{ij}} \phi_j^\dagger \right) + \sum_{j \leqslant N < i} \left(\phi_i \omega_{ij} \frac{\partial E}{\partial P_{ij}} \right) \phi_j^\dagger \\
&= \phi \bar{J} \phi^\dagger + \phi \psi^\dagger + \psi \phi^\dagger
\end{aligned}
$$

where ψ is $N_b \times N$ with $\psi_j = \sum_{N<i} \phi_i \left(\omega_{ij} \frac{\partial E}{\partial P_{ij}} \right)$ and $\bar{J}$ is the $N \times N$ (Hermitian) matrix given by

$$\bar{J}_{ij} = \omega_{ij} \left(\frac{\partial E}{\partial P_{ij}} - \frac{\chi}{\zeta} \delta_{ij} \right) = \omega_{ij} \left(\phi_i^{\dagger} \frac{\partial E}{\partial P} \phi_j - \frac{\chi}{\zeta} \delta_{ij} \right)$$

requiring only matrix elements on ϕ. This gives an orbital form of (11),

$$\frac{\partial E}{\partial V(r)} = \sum_{i,j=1}^{N} \phi_i(r) \bar{J}_{ij} \phi_j^*(r) + \sum_{i=1}^{N} \left(\phi_i(r) \psi_i^*(r) + \psi_i(r) \phi_i^*(r) \right) \qquad (13)$$

which is similar to the equation derived in numerous sources in the OEP literature [13, 14], with a modification (the $\bar{J}$ term) to accommodate the finite temperature regime.

If full diagonalization were possible, then one need only evaluate $\frac{\partial E}{\partial P_{ij}}$, which by (7) is

$$\frac{\partial E}{\partial P_{ij}} = \phi_i^{\dagger} \frac{\partial E}{\partial P} \phi_j = \int \int \phi_i(r) \left[\frac{\partial E_{\text{HXC}}}{\partial P(r,r')} - \delta(r-r')V(r) \right] \phi_j(r') \, d^3r \, d^3r',$$

and carry out the sums. Alternatively, for any $j \leqslant N < i$, we have

$$(\varepsilon_i - \varepsilon_j)\omega_{ij} = -\omega_j$$

$$(\varepsilon_i - \varepsilon_j)\phi_i \omega_{ij} \frac{\partial E}{\partial P_{ij}} = -\omega_j \phi_i \frac{\partial E}{\partial P_{ij}}$$

$$\sum_{N<i} \varepsilon_i \phi_i \omega_{ij} \frac{\partial E}{\partial P_{ij}} - \varepsilon_j \sum_{N<i} \phi_i \omega_{ij} \frac{\partial E}{\partial P_{ij}} = -\sum_{N<i} \phi_i \frac{\partial E}{\partial P_{ij}} \omega_j$$

which can be expressed in basis-free terms as

$$(H - \varepsilon_j I)\psi_j = -(I - \phi\phi^{\dagger}) \frac{\partial E}{\partial P} \phi_j \omega_j. \qquad (14)$$

The left- and right-hand sides of (14) are orthogonal to ϕ, by construction, as is each ψ_j. Thus an iterative method (e.g. conjugate gradient or MINRES [19, 20]) can be used to solve for ψ while preserving the constraint $\phi^{\dagger}\psi = 0$.

We digress momentarily to recover the results of [14] in the zero temperature limit. Assuming that $\lambda_N < \mu < \lambda_{N+1}$,

$$\omega_i = \begin{cases} 1 & i \leqslant N \\ 0 & i > N \end{cases} \qquad \text{and} \qquad \omega_{ij} = \begin{cases} \dfrac{1}{\varepsilon_i - \varepsilon_j} & i \leqslant N < j \\ \dfrac{1}{\varepsilon_j - \varepsilon_i} & j \leqslant N < i \\ 0 & \text{else.} \end{cases}$$

$\bar{J} = 0$, $\chi = \zeta = 0$ and ψ is given by,

$$\psi_j(r) = \sum_{i>N} \frac{\phi_i(r)}{\varepsilon_i - \varepsilon_j} \frac{\partial E}{\partial P_{ij}}$$

$$= \sum_{i>N} \frac{\phi_i(r)}{\varepsilon_i - \varepsilon_j} \int \int \phi_i(r) \left[\frac{\partial E_{\text{HXC}}}{\partial P(r,r')} - \delta(r-r')V(r) \right] \phi_j(r') \, d^3r \, d^3r'.$$

The only non-trivial difference between this result and that of [14] is the sum defining $\psi_j(r)$, which, in (3) of [14], ranges over all $i \neq j$ instead of $i > N$. Such ψ_i are the unnormalized orbital shifts of the perturbed Hamiltonian $H + \delta H$ and are of some interest. However, $\frac{\partial E}{\partial P_{ij}} = \frac{\partial E^*}{\partial P_{ji}}$ ensures that this difference vanishes in the sum in (13).

Finally, we note that $\omega_{ii} \sim \frac{1}{4}\beta$ if $\varepsilon_i \sim \mu$, which suggests that in the low temperature limit numerical instability might result if any of the eigenvalues of H are close to the Fermi level.

6. 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 ψ 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 LDA and GGA, E_{HXC} is an explicit functional of the electronic density, which is the diagonal of the density matrix P in a position representation. In this case, $\frac{\partial E_{HXC}}{\partial P}$ 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.

In the case of exact exchange (EXX) calculations, we neglect correlation and take E_{HXC} to be the sum of the Hartree and exchange energies. The evaluation of E_{HXC} and $\frac{\partial E_{HXC}}{\partial P}$ involves the same exchange integrals as required in a Hartree–Fock calculation, but they only need to be performed once per gradient evaluation. We follow the approach of Gygi and Baldereschi to evaluate the exchange integrals using a plane wave basis [21].

Our test system consists of a two-atom unit cell of silicon in the diamond structure. We used a 20 Ryd 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).
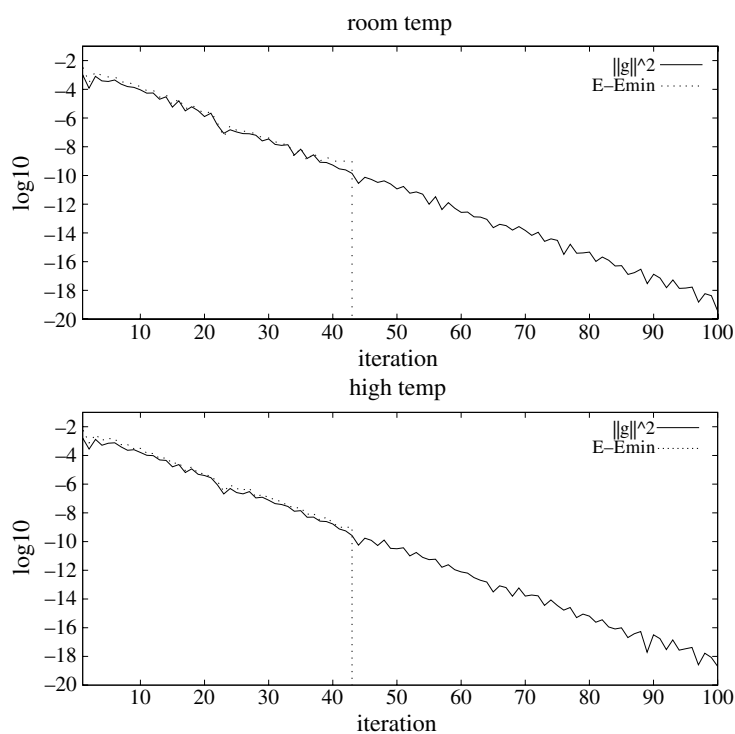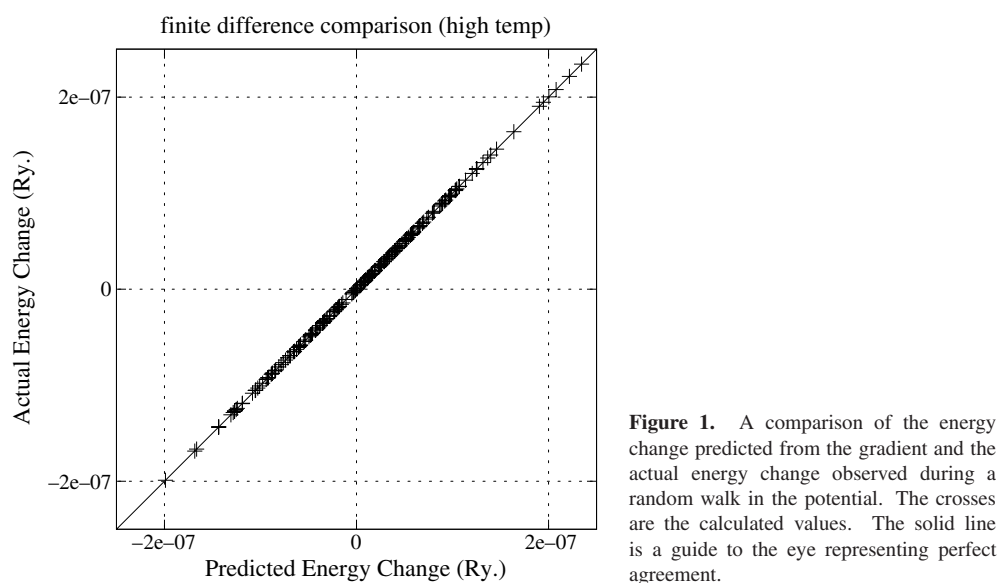
In order to test the correctness of our gradient, we applied the finite difference approach to the EXX energy functional. During each of a series of steps, the value of V at each point on a real-space grid was varied by a small (o(10^{-4})) random perturbation $\Delta V(r)$. During this random walk, the energy and the gradient were evaluated at each step. A linear approximation to the change in energy during each step is given by

$$\Delta E \approx \int \frac{\partial E}{\partial V(r)} \Delta V(r) \, dr. \tag{15}$$

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 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 minimization step replaces the standard self-consistency iteration in our code with the gradient calculation and potential update replacing the conventional density mixing. The convergence of the energy of our test system during this process is shown in figures 2 and 3. In figure 2, we used our approach to perform a LDA calculation. The errors in the energy during the iterative minimization are computed by comparison 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. This shows that our minimization

finite difference comparison (high temp)

Figure 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 2. The error in the LDA 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 (13).

approach is reaching the minimum of the energy functional. In figure 3, we performed an EXX calculation. The errors in the energy were approximated by comparison to the converged result.
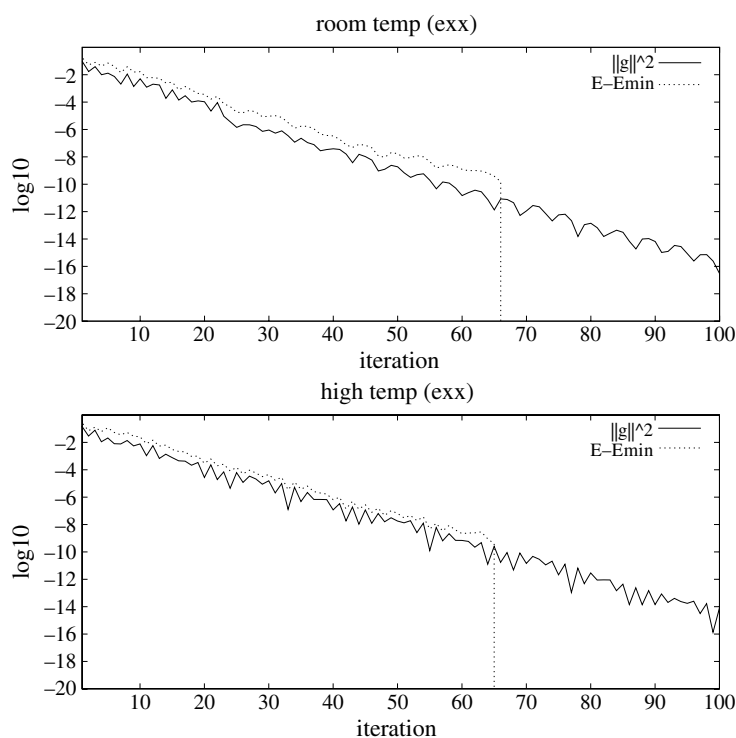
Figure 3. The convergence of the *exact exchange* energy, as well as the square-norm of gradient, during iterative minimization on a log scale. The top plot was run at room temperature (25.67 meV) and the bottom at high temperature (1 eV). $g = \frac{\partial E}{\partial V(r)}$ from (13).

In both the LDA and EXX cases, we found that 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} Ryd) can be obtained easily.

7. 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.

Acknowledgments

References

[1] Kohn W L and Sham L J 1965 *Phys. Rev.* **140** A1133
[2] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[3] Talman J D and Shadwick W F 1976 *Phys. Rev.* A **14** 36
[4] Sahni V, Gruenebaum J and Perdew J P 1982 *Phys. Rev.* B **26** 4371
[5] Engel E and Vosko S H 1993 *Phys. Rev.* A **47** 2800
    Engel E and Vosko S H 1994 *Phys. Rev.* B **50** 10498
[6] Görling A and Levy M 1994 *Phys. Rev.* A **50** 196
[7] Kotani T 1995 *Phys. Rev. Lett.* **74** 2989
[8] Städele M, Majewski J A, Vogl P and Görling A 1997 *Phys. Rev. Lett.* **79** 2089
    Städele M, Moukara M, Majewski J A, Vogl P and Görling A 1999 *Phys. Rev.* B **59** 10031
[9] Görling A 1999 *Phys. Rev. Lett.* **83** 5459
[10] Kümmel S, Kronik L and Perdew J P 2004 *Phys. Rev. Lett.* **93** 213002
[11] Perdew J P, Ruzsinszky A, Tao J, Staroverov V N, Scuseria G E and Csonka G I 2005 *J. Chem. Phys.* **123** 62201
[12] Yang W and Wu Q 2002 *Phys. Rev. Lett.* **89** 143002
[13] Hyman R, Stiles M and Zangwill A 2000 *Phys. Rev.* B **62** 15521
[14] Kümmel S and Perdew J P 2003 *Phys. Rev. Lett.* **90** 43004
[15] Kümmel S and Perdew J P 2003 *Phys. Rev.* B **68** 35103
[16] Desjarlais M P, Kress J D and Collins L A 2002 *Phys. Rev.* E **66** 25401
[17] Desjarlais M P 2003 *Phys. Rev.* B **68** 64204
[18] Horn R A and Johnson C R 1991 *Topics in Matrix Analysis* (Cambridge: Cambridge University Press)
[19] Paige C C and Saunders M A 1975 *SIAM J. Numer. Anal.* **12** 617
[20] Golub G H and Loan C F V 1989 *Matrix Computations* 2nd edn (Baltimore, MD: Johns Hopkins University Press)
[21] Gygi F and Baldereschi A 1986 *Phys. Rev.* B **34** 4405