PHYSICAL REVIEW B **84**, 165122 (2011)

Direct minimization of the optimized effective problem based on efficient finite differences

Chen Huang

Department of Physics, Princeton University, Princeton, New Jersey 08544, USA

Emily A. Carter[*]

Department of Mechanical and Aerospace Engineering, Program in Applied and Computational Mathematics, and the Andlinger Center for Energy and the Environment, Princeton University, Princeton, New Jersey 08544-5263, USA

(Received 22 July 2011; revised manuscript received 30 September 2011; published 24 October 2011)

To avoid the difficult-to-solve optimized effective potential (OEP) integral equation, we introduce an efficient direct minimization scheme for performing OEP calculations within Kohn–Sham density functional theory (KS-DFT). We reformulated the functional derivative of the total energy with respect to the KS effective potential in terms of efficient finite differences. Our method only uses the orbitals involved in the construction of the KS exchange-correlation functionals. We demonstrate our scheme by performing exact-exchange OEP for sodium clusters, in which only occupied KS orbitals are needed to obtain the OEP. Our efficient direct minimization scheme should aid future development of orbital-dependent density functionals and render OEP to be a practical choice for various applications.

DOI: 10.1103/PhysRevB.84.165122 PACS number(s): 31.15.E−

I. INTRODUCTION

Kohn–Sham density functional theory (KS-DFT) is formally an exact theory[1] based on the Hohenberg–Kohn theorems.[2] In practice, the accuracy of KS-DFT is severely limited by the approximations used in the exchange-correlation (XC) functionals. To improve the accuracy of KS-DFT, we may think of climbing a so-called Jacob's ladder[3] for XC functionals, from the local density approximation (LDA), the semilocal generalized gradient approximation (GGA), to meta-GGAs that depend on the Laplacian of the density and KS kinetic energy densities, to orbital-dependent XC functionals[4] (ODXCFs), that explicitly depend on the KS orbitals, instead of only on the density. One straightforward example of ODXCFs is the EXact-eXchange (EXX) functional, i.e. Hartree–Fock exchange with KS orbitals as input. It has been demonstrated that ODXCFs are critical to describing many physical phenomena, such as the discontinuity of the XC potential during electron transfer,[5] electrical response of molecules,[6] and spin dynamics in noncollinear magnets,[7] to name a few. For all these cases, KS-DFT using LDA or GGA XC often fails completely. Even though ODXCFs are highly desirable for performing reliable KS-DFT calculations, we have to find the effective XC potential $v_{xc}(\vec{r})$ associated with ODXCFs by solving the cumbersome optimized effective potential (OEP) equations. An early formulation of such an OEP equation[8] is a Fredholm integral equation of the first kind that is difficult to solve[9] and engenders huge computational costs in constructing the KS Green's function, which in principle requires calculating not only the occupied but also all virtual KS orbitals. By introducing orbital shifts, the OEP equation can be reformulated as a set of two coupled integral equations for $v_{xc}(\vec{r})$ and orbital shifts.[10] One advantage of this reformulation is that only the orbitals used in the definition of XC functionals are needed in the OEP calculations. However, the resulting equations remain challenging to solve. More critically, the equation for $v_{xc}(\vec{r})$ is numerically unstable for finite systems.[11] To avoid such numerical instabilities, Kümmel and Perdew[11] solved for $v_{xc}(\vec{r})$ iteratively by correcting it with a scaled residual. It can be shown that this residual is exactly the gradient of the total energy with respect to the KS effective potential $v_S(\vec{r})$ at that iteration, even though they did not claim this. Therefore, their iterative method searches for the $v_{xc}(\vec{r})$ in the steepest decent direction and is inefficient when far away from the solution.

To avoid solving OEP equations, a general and promising approach is to perform a direct minimization of the total energy with respect to $v_S(\vec{r})$.[12–15] Hyman *et al.*[16] demonstrated quite some time ago a simple, general, and efficient direct minimization scheme in which the gradient of total energy with respect to the KS effective potential was obtained by solving the orbital shift equations (they called these quantities Lagrange functions instead of orbital shifts), which are nearly the same as the orbital shift equations derived by Kümmel and Perdew.[11] In this paper, instead of solving orbital shift equations, we reformulate the gradient of the total energy in terms of easy-to-compute finite differences, completely eliminating the use of perturbation theory present in Yang and Wu's direct minimization scheme.[15] Our work is an extension of Gonze's work[17] for treating ODXCFs.

II. THEORY

We start with the collinear spin-polarized case. The KS equations are

$$\hat{H}_{KS,\sigma}\phi_{K,\sigma}(\vec{r}) = \varepsilon_{K,\sigma}\phi_{K,\sigma}(\vec{r}). \tag{1}$$

Here, σ is the spin index, K is an orbital index, $\hat{H}_{KS,\sigma} = -\frac{1}{2}\nabla^2 + v_{s,\sigma}(\vec{r})$ is the KS one-electron Hamiltonian, $\phi_{K,\sigma}(\vec{r})$ and $\varepsilon_{K,\sigma}$ are the eigenfunction and eigenvalue of Eq. (1), and therefore functionals of $v_{s,\sigma}(\vec{r})$. We define the orbital energy as a functional of $v_{s,\sigma}(\vec{r})$:

$$E_{K,\sigma}[v_{s,\sigma}] = \langle\phi_{K,\sigma}[v_{s,\sigma}]| -\tfrac{1}{2}\nabla^2 + v_{s,\sigma}(\vec{r})|\phi_{K,\sigma}[v_{s,\sigma}]\rangle. \tag{2}$$

The orbital density is defined as $\rho_{K,\sigma}(\vec{r}) = |\phi_{K,\sigma}(\vec{r})|^2$. The Hellmann–Feynman theorem[18] allows us to write the functional derivative of $E_{K,\sigma}[v_{s,\sigma}]$ as

$$\frac{\delta E_{K,\sigma}[v_{s,\sigma}]}{\delta v_{s,\sigma}(\vec{r})} = \rho_{K,\sigma}(\vec{r}). \tag{3}$$

Using Eq. (3), we reformulate the gradient of the total energy with respect to $v_{s,\sigma}(\vec{r})$ in terms of finite differences as follows:

$$\begin{aligned}\frac{\delta E_{\text{tot}}}{\delta v_{s,\sigma}(\vec{r})} &= \sum_K \int \frac{\delta E_{\text{tot}}}{\delta \rho_{K,\sigma}(\vec{r}')} \frac{\delta \rho_{K,\sigma}(\vec{r}')}{\delta v_{s,\sigma}(\vec{r})} dr'^3 \\ &= \sum_K \int \frac{\delta E_{\text{tot}}}{\delta \rho_{K,\sigma}(\vec{r}')} \frac{\delta^2 E_{K,\sigma}[v_{s,\sigma}]}{\delta v_{s,\sigma}(\vec{r}) \delta v_{s,\sigma}(\vec{r}')} dr'^3 \\ &= \sum_K \int \frac{\delta E_{\text{tot}}}{\delta \rho_{K,\sigma}(\vec{r}')} \frac{\delta \rho_{K,\sigma}(\vec{r})}{\delta v_{s,\sigma}(\vec{r}')} dr'^3 \\ &\approx \sum_K \frac{1}{\lambda_{K,\sigma}} \{\rho_{K,\sigma}[v_{s,\sigma} + \lambda_{K,\sigma} \Delta v_{s,K,\sigma}](\vec{r}) \\ &\quad - \rho_{K,\sigma}[v_{s,\sigma}](\vec{r})\},\end{aligned} \tag{4}$$

where $\lambda_{K,\sigma}$ is a small step size for the finite difference and $\Delta v_{s,K,\sigma}(\vec{r}) = \delta E_{\text{tot}}/\delta \rho_{K,\sigma}(\vec{r})$. From the third line to the fourth line in Eq. (4), we have used the definition of the functional derivative. Here, $\rho_{K,\sigma}[v_{s,\sigma} + \lambda_{K,\sigma} \Delta v_{s,K,\sigma}]$ means that a new electron density is calculated for a slightly changed potential: $v_{s,\sigma} + \lambda_{K,\sigma} \Delta v_{s,K,\sigma}$. At the minimum of E_{tot}, $\Delta v_{s,K,\sigma}(\vec{r})$ is equal to the chemical potential, which is a constant in space, and therefore, Eq. (4) becomes equal to zero.

When computing $\delta E_{\text{tot}}/\delta \rho_{K,\sigma}$, one complication that arises is the need to evaluate $\delta T_S/\delta \rho_{K,\sigma}$; all other components in E_{tot} are explicit functionals of $\rho_{K,\sigma}(\vec{r}')$ that are straightforwardly calculated. Because the KS orbitals are solved for by minimizing the nonself-consistent (NSC) energy functional[17] $E^{\text{NSC}} = \sum_\sigma \sum_{K=\text{occ}} \langle \phi_{K,\sigma} | -\frac{1}{2}\nabla^2 + v_{s,\sigma} | \phi_{K,\sigma} \rangle$, then the functional derivative of the KS kinetic energy functional is simply $\delta T_S/\delta \rho_{K,\sigma}(\vec{r}) = \delta E^{\text{NSC}}/\delta \rho_{K,\sigma} - v_{s,\sigma}(\vec{r}) = \varepsilon_{K,\sigma} - v_{s,\sigma}(\vec{r})$. Now, with the gradient in Eq. (4) in hand, efficient OEP calculations can be performed.

III. NUMERICAL METHODS

We demonstrate the above direct minimization method by performing non-spin-polarized exchange-only KS-DFT calculations for the Na_4 and Na_8 clusters. Our method is implemented in the ABINIT[19] plane-wave-based KS-DFT code. The EXX functional is used

$$E_X = -\frac{1}{4} \sum_{m=\text{occ}} \sum_{n=\text{occ}} f_n f_m \times \iint \frac{\phi_m^*(\vec{r})\phi_m(\vec{r}')\phi_n(\vec{r})\phi_n^*(\vec{r}')}{|\vec{r} - \vec{r}'|} dr^3 dr'^3,$$

where f_n is the occupation number for orbital n. To evaluate E_X, we compute the integral over $\vec{r}'$ with the Poisson solver[20] in ABINIT[19] and then the integral over $\vec{r}$ is obtained by direct summation in the real space. The total energy E_{tot} is minimized using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton method.[21] To calculate the new orbital densities for these slightly changed potentials $v_{s,\sigma}(\vec{r}) + \lambda_{K,\sigma} \Delta v_{s,K,\sigma}(\vec{r})$ in Eq. (4), we use the spectrum-folding method,[22] in which the new $\phi_{K,\sigma}^{\text{new}}$ is obtained by minimizing the square residual $\min_{\phi_{K,\sigma}} \langle \phi_{K,\sigma} | (\hat{H}_{KS} + \lambda_{K,\sigma} \Delta v_{s,K,\sigma}(\vec{r}) - \varepsilon_{K,\sigma})^2 | \phi_{K,\sigma} \rangle$. The spectrum-folding method delivers the eigenfunction that is closest to $\varepsilon_{K,\sigma}$, which renders the calculations of the new $\phi_{K,\sigma}^{\text{new}}$ to be highly efficient. A conjugate gradient method is employed to perform the spectrum folding. The step size $\lambda_{K,\sigma}$ is determined by setting the quantity $\langle \phi_{K,\sigma} | \lambda \Delta v_{s,K,\sigma}(\vec{r}) | \phi_{K,\sigma} \rangle$ equal to 3 mHa. This quantity is in fact the first order correction to $\varepsilon_{K,\sigma}$ with respect to the change of potential. This self-adjusted step size $\lambda_{K,\sigma}$ guarantees a moderate change of $\phi_{K,\sigma}$.

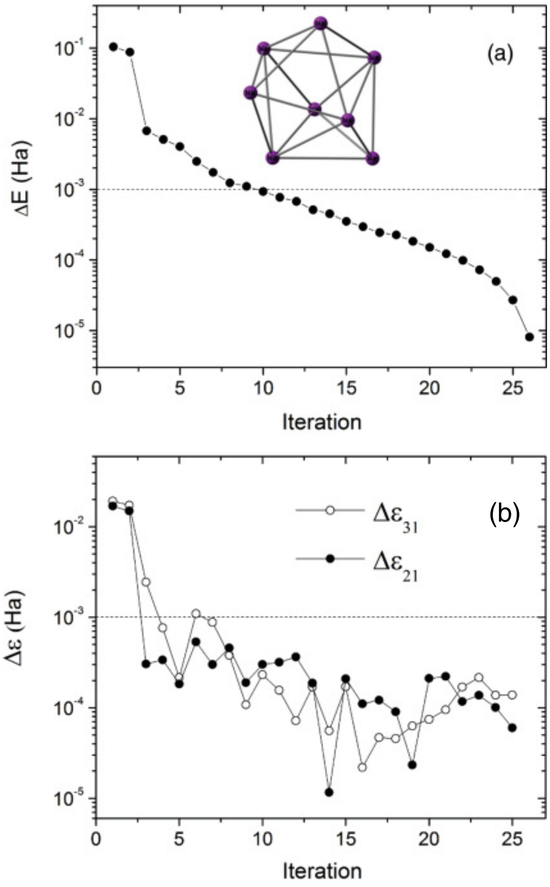

FIG. 1. (a) The convergence of the total energy, $\Delta E = E_{\text{tot}} - E_{\text{tot,final}}$, vs BFGS iterations for the Na_8 cluster. Inset shows the geometry of the cluster. (b) The convergence of eigenvalue differences $\Delta\varepsilon_{21} = (\varepsilon_2 - \varepsilon_1) - (\varepsilon_2 - \varepsilon_1)_{\text{final}}$ (solid circle) and $\Delta\varepsilon_{31} = (\varepsilon_3 - \varepsilon_1) - (\varepsilon_3 - \varepsilon_1)_{\text{final}}$ (hollow circle) vs BFGS iterations for the Na_8 cluster.

To compare our results with previous calculations, the Na local pseudopotential and the geometries for Na_4 and Na_8 clusters in Ref. 11 are used. The kinetic energy cutoff for the size of plane wave basis set is 600 eV. For Na_8 (see the inset of Fig. 1 for the structure), the simulation box is 30 × 30 × 30 Å. For Na_4, the cluster is a planar rhombus and is placed on the x-z plane (see the inset in Fig. 2 for the structure). The simulation box is 40 × 20 × 40 Å.

IV. RESULTS AND DISCUSSION

To demonstrate the feasibility and accuracy of our direct minimization method, we compare the total energies and eigenvalues of the Na_4 and Na_8 clusters to the previous calculations by Kümmel *et al.*[11] (Table I). Since $v_s(\vec{r})$ is only determined up to a constant shift, we report the eigenvalue differences $\varepsilon_4 - \varepsilon_1$, $\varepsilon_3 - \varepsilon_1$, and $\varepsilon_2 - \varepsilon_1$, where ε_i, $i = 1,\ldots,4$ are the eigenvalues of the four lowest eigenstates (see Table I). All the eigenvalue differences and total energies agree very well with Kümmel *et al.*'s[11] results, with a maximum difference of 0.3 mHa. The virial theorem,[23] $E_X + \int dr^3 \rho(\vec{r})\vec{r}\cdot\vec{\nabla}v_X(\vec{r}) = 0$, is satisfied with a relative error (defined as $(E_X + \int dr^3\rho(\vec{r})\vec{r}\cdot\vec{\nabla}v_X(\vec{r}))/E_X$) of 3% and 0.1% for the Na_4 and Na_8, respectively. The errors are mainly due to the slightly inaccurate asymptotic behavior of XC potential away from the clusters, which causes Na_4 to have a relatively larger error, while the error of the larger cluster Na_8 is rather small. However, this slightly inaccurate asymptotic behavior has a very small effect on the total energies and eigenvalue differences as demonstrated here. For the Na_4 cluster, the total energy and the eigenvalue differences converge to within 1 mHa after only six and four iterations, respectively. For the Na_8 cluster, the total energy and eigenvalue differences converge quickly to within 1 mHa after only 10 iterations [see Fig. 1(a)] and seven iterations [see Fig. 1(b)], respectively. This swift convergence does not require a very good initial guess for $v_s(\vec{r})$. The $v_s(\vec{r})$ is initialized based on the total electron density from a superposition of Gaussian-like atomic electron densities, and LDA XC functional is used for initializing $v_{xc}(\vec{r})$. Therefore, our direct minimization is numerically more appealing than Kümmel *et al.*'s approach,[11] in which a good initial guess for $v_s(\vec{r})$ obtained using the Krieger–Li–Iafrate (KLI) approximation was used to accelerate the OEP convergence. However, the KLI approximation has exhibited convergence problems in some cases.[6]

In Fig. 2, our $v_x(\vec{r})$ goes to zero in all directions. By contrast, Kümmel *et al.*[11] found a nonvanishing asymptotic behavior of $v_x(\vec{r})$ in the nodal plane of the p-like highest occupied orbital even far away from the cluster. The reason for this discrepancy is twofold. (i) In the direct minimization scheme, the $v_x(\vec{r})$ is obtained by minimizing the total energy, which is not sensitive to the asymptotic behavior of $v_x(\vec{r})$ far from the cluster where the wave function decays to zero. Therefore, the program stops searching for the optimal $v_x(\vec{r})$ before such distant nonuniform asymptotic behavior of $v_x(\vec{r})$ develops. (ii) The nonuniform asymptotic behavior of $v_x(\vec{r})$ only exists in the nodal plane of measure zero, which in turn produces the so-called barrier-well structure in the near-asymptotic region.[24] Far away from the cluster, this barrier-well structure should become thinner such that its contribution to the total energy lessens. Nevertheless, near the cluster, we indeed observe this barrier-well structure (see black arrow in Fig. 2). In our numerical tests, we find that quantities related to the ground state, such as the total energy and eigenvalues of occupied KS orbitals, can be obtained very accurately.

TABLE I. Comparison of the OEP-EXX total energies E_{tot} and eigenvalue differences obtained from our direct minimization and Kümmel *et al.*[11] All results are in Hartrees.

Na_4	E_{tot}	$\varepsilon_2 - \varepsilon_1$		
Direct	−0.7529	0.0359		
Kümmel	−0.7531	0.0361		
Na_8	E_{tot}	$\varepsilon_2 - \varepsilon_1$	$\varepsilon_3 - \varepsilon_1$	$\varepsilon_4 - \varepsilon_1$
Direct	−1.5282	0.0361	0.0492	0.0494
Kümmel	−1.5285	0.0362	0.0495	0.0495

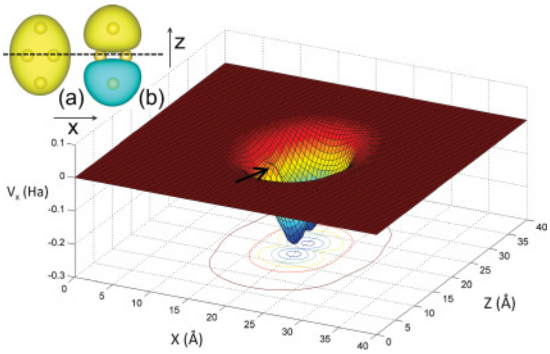

FIG. 2. (Color online) Contour plot of the exchange potential $v_x(\vec{r})$ of the Na_4 cluster in the cluster plane (x-z plane). (Top left) Two occupied orbitals calculated by KS-DFT-EXX. (a) The lowest energy orbital with s character. (b) The highest occupied orbital, with p character. Sodium atoms are shown as yellow balls. The nodal plane of the p-like orbital is marked with black dashed line. (Main plot) The black arrow points out the barrier-well structure, which disappears away from the cluster.

V. EXTENSION TO CURRENT-SPIN DENSITY FUNCTIONAL THEORY

Finally, we show that the direct minimization method introduced in this paper can be generalized to many other KS systems. Here, we extend it to the current-spin density functional theory (CSDFT).[25] This is an important extension. For example, use of the OEP-EXX has been found to be critical for producing spin dynamics in SDFT and CSDFT calculations that are absent when LDA and GGA are used.[7] In CSDFT, the total energy, subject to an external scalar potential $v_{\mathrm{ext}}(\vec{r})$, a magnetic field $\vec{B}_{\mathrm{ext}}(\vec{r})$, and a vector potential $\vec{A}_{\mathrm{ext}}(\vec{r})$, in the KS form is[25]

$$\begin{aligned} E_{\mathrm{tot}}[\rho,\vec{m},\vec{j}_p] &= T_S[\rho,\vec{m},\vec{j}_p] + E_{XC}[\rho,\vec{m},\vec{j}_p] + J[\rho] \\ &\quad + \int v_{\mathrm{ext}}(\vec{r})\rho(\vec{r})dr^3 + \int \vec{m}(\vec{r})\cdot\vec{B}_{\mathrm{ext}}(\vec{r})dr^3 \\ &\quad + \frac{1}{c}\int \vec{j}_p(\vec{r})\cdot\vec{A}_{\mathrm{ext}}(\vec{r})dr^3 + \frac{1}{2c^2}\int \rho(\vec{r})A_{\mathrm{ext}}(\vec{r})^2 dr^3. \end{aligned} \tag{5}$$

The KS orbitals are two-component spinors $\Phi_K(\vec{r}) = \binom{\phi_{K,\alpha}(\vec{r})}{\phi_{K,\beta}(\vec{r})}$ for which K is the orbital index. The magnetization is $\vec{m}(\vec{r}) = \mu_B \sum_{K=occ} \Phi_K^+(\vec{r})\vec{\sigma}\Phi_K(\vec{r})$ with $\vec{\sigma}$ being the Pauli matrix. The current is $\vec{j}_p = \frac{1}{2i}\sum_{K=occ}\{\Phi_K^+\vec{\nabla}\Phi_K - (\vec{\nabla}\Phi_K^+)\Phi_K\}$. Here,

$T_S[\rho,\vec{m},\vec{j}_p]$, $E_{XC}[\rho,\vec{m},\vec{j}_p]$, and $J[\rho]$ are the KS kinetic energy, XC functional, and classical Hartree electron repulsion energy, respectively.[25] Then the CSDFT KS equation is

$$\hat{H}_{KS}\Phi_K = \varepsilon_K \Phi_K, \tag{6}$$

with

$$\hat{H}_{KS} = -\frac{1}{2}\vec{\nabla}^2 + v_s(\vec{r}) + \mu_B\vec{\sigma}\cdot\vec{B}_s(\vec{r}) + \frac{1}{2ic}[\vec{A}_s(\vec{r})\cdot\vec{\nabla} + \vec{\nabla}\cdot\vec{A}_s(\vec{r})].$$

The KS effective fields are defined as

$$v_s(\vec{r}) = v_{\text{ext}}(\vec{r}) + \frac{\delta J[\rho]}{\delta\rho(\vec{r})} + \left.\frac{\delta E_{XC}[\rho,m,\vec{j}_p]}{\delta\rho(\vec{r})}\right|_{\vec{m},\vec{j}_p} + \frac{1}{2c^2}A_{\text{ext}}(\vec{r})^2,$$

$$\vec{B}_s(\vec{r}) = \vec{B}_{\text{ext}}(\vec{r}) + \left.\frac{\delta E_{XC}[\rho,m,\vec{j}_p]}{\delta\vec{m}(\vec{r})}\right|_{\rho,\vec{j}_p},$$

$$\vec{A}_s(\vec{r}) = \vec{A}_{\text{ext}}(\vec{r}) + c\left.\frac{\delta E_{XC}[\rho,m,\vec{j}_p]}{\delta\vec{j}_p(\vec{r})}\right|_{\rho,\vec{m}}.$$

Note that the total energy is therefore an implicit functional of $v_s(\vec{r})$, $\vec{B}_s(\vec{r})$, and $\vec{A}_s(\vec{r})$ via the density matrix $\rho_{\alpha\beta} = \sum_{K=occ}\phi_{K,\alpha}\phi^*_{K,\beta}$ and current $\vec{j}_p$. Following a similar procedure, we define the orbital-related quantities: orbital energy $E_K[v_s, B_s, A_s]$, orbital density matrix $\rho_{\alpha\beta,K}(\vec{r})$, and orbital current $\vec{j}_{p,K}(\vec{r})$

$$E_K[v_s,B_s,A_s] = \langle\Phi_K[v_s,B_s,A_s]|\widehat{H}_{KS}|\Phi_K[v_s,B_s,A_s]\rangle, \tag{7}$$

$$\rho_{\alpha\beta,K}(\vec{r}) = \phi_{K,\alpha}(\vec{r})\phi^*_{K,\beta}(\vec{r}), \tag{8}$$

$$\vec{j}_{p,K}(\vec{r}) = \frac{1}{2i}\{\Phi^+_K(\vec{r})\vec{\nabla}\Phi_K(\vec{r}) - [\vec{\nabla}\Phi^+_K(\vec{r})]\Phi_K(\vec{r})\}. \tag{9}$$

The Kth orbital density is $\rho_K(\vec{r}) = \sum_\sigma \rho_{\sigma\sigma,K}(\vec{r})$. Making use of the Hellmann–Feynman theorem,[18] we perform functional derivatives of $E_K[v_s, B_s, A_s]$ with respect to these KS effective fields, and define new variables $\vec{a}_K(\vec{r})$, $\vec{b}_K(\vec{r})$, and $c_K(\vec{r})$

$$\frac{\delta E_K}{\delta\vec{A}_s(\vec{r})} = \frac{1}{c}\vec{j}_{p,K}(\vec{r}) \equiv \vec{a}_K(\vec{r}), \tag{10}$$

$$\frac{\delta E_K}{\delta\vec{B}_s(\vec{r})} = \mu_B\Phi^+_K\vec{\sigma}\Phi_K \equiv \vec{b}_K(\vec{r}), \tag{11}$$

$$\frac{\delta E_K}{\delta v_s(\vec{r})} = \rho_K(\vec{r}) \equiv c_K(\vec{r}). \tag{12}$$

The gradient of E_{tot} with respect to $A_{s,1}(\vec{r})$ (the x component of $\vec{A}_s(\vec{r})$, and similar procedures for $A_{s,2}(\vec{r})$ and $A_{s,3}(\vec{r})$), by the chain rule, is

$$\begin{aligned}\frac{\delta E_{\text{tot}}}{\delta A_{s,1}(\vec{r})} &= \sum_K\sum_{j=1}^3\int\frac{\delta E_{\text{tot}}}{\delta a_{K,j}(\vec{r}')}\frac{\delta a_{K,j}(\vec{r}')}{\delta A_{s,1}(\vec{r})}dr'^3 \\ &\quad + \sum_K\sum_{j=1}^3\int\frac{\delta E_{\text{tot}}}{\delta b_{K,j}(\vec{r}')}\frac{\delta b_{K,j}(\vec{r}')}{\delta A_{s,1}(\vec{r})}dr'^3 \\ &\quad + \sum_K\int\frac{\delta E_{\text{tot}}}{\delta c_K(\vec{r}')}\frac{\delta c_K(\vec{r}')}{\delta A_{s,1}(\vec{r})}dr'^3.\end{aligned} \tag{13}$$

The derivative of E_{tot} with respect to $\vec{a}_K(\vec{r})$, $\vec{b}_K(\vec{r})$, and $c_K(\vec{r})$ can be calculated by the chain rule after expressing $\rho_{\alpha\beta,K}(\vec{r})$ and $\vec{j}_{p,K}(\vec{r})$ in terms of $\vec{a}_K(\vec{r})$, $\vec{b}_K(\vec{r})$, and $c_K(\vec{r})$ with the help of Eqs. (10)–(12). The index K only runs over all the orbitals used in the construction in the XC functionals. The first integral in Eq. (13) can be reformulated with finite differences, with the help of Eq. (10)

$$\begin{aligned}&\int\frac{\delta E_{\text{tot}}}{\delta a_{K,j}(\vec{r}')}\frac{\delta a_{K,j}(\vec{r}')}{\delta A_{s,1}(\vec{r})}dr'^3 \\ &= \int\frac{\delta E_{\text{tot}}}{\delta a_{K,j}(\vec{r}')}\frac{\delta^2 E_K}{\delta A_{s,1}(\vec{r})\delta A_{s,j}(\vec{r}')}dr'^3 \\ &= \int\frac{\delta E_{\text{tot}}}{\delta a_{K,j}(\vec{r}')}\frac{\delta a_{K,1}(\vec{r})}{\delta A_{s,j}(\vec{r}')}dr'^3 \\ &\approx \frac{1}{\lambda}\left(a_{K,1}\left[A_{s,j}\to A_{s,j}+\lambda\frac{\delta E_{\text{tot}}}{\delta a_{K,j}}\right](\vec{r}) - a_{K,1}(\vec{r})\right).\end{aligned} \tag{14}$$

Above, $a_{K,1}[A_{s,j}\to A_{s,j}+\lambda\frac{\delta E_{\text{tot}}}{\delta a_{K,j}}](\vec{r})$ means that a new $a_{K,1}(\vec{r})$ is calculated using Eq. (6) with $A_{s,\text{j}}$ being replaced by $A_{s,j}+\lambda\frac{\delta E_{\text{tot}}}{\delta a_{K,j}}$, and the other components of effective fields are unchanged. Again, λ is a small finite difference. The same notation is used throughout this work.

Similarly, using Eq. (11), the second integral in Eq. (13) is reformulated to

$$\begin{aligned}&\int\frac{\delta E_{\text{tot}}}{\delta b_{K,j}(\vec{r}')}\frac{\delta b_{K,j}(\vec{r}')}{\delta A_{s,1}(\vec{r})}dr'^3 \\ &\approx \frac{1}{\lambda}\left\{a_{K,1}\left[B_{s,j}\to B_{s,j}+\lambda\frac{\delta E_{\text{tot}}}{\delta b_{K,j}(\vec{r})}\right](\vec{r}) - a_{K,1}(\vec{r})\right\}.\end{aligned} \tag{15}$$

Using Eq. (12), the third integral in Eq. (13) becomes

$$\begin{aligned}&\int\frac{\delta E_{\text{tot}}}{\delta c_K(\vec{r}')}\frac{\delta c_K(\vec{r}')}{\delta A_{s,1}(\vec{r})}dr'^3 \\ &\approx \frac{1}{\lambda}\left\{a_{K,1}\left[v_s\to v_s+\lambda\frac{\delta E_{\text{tot}}}{\delta c_K(\vec{r})}\right](\vec{r}) - a_{K,1}(\vec{r})\right\}.\end{aligned} \tag{16}$$

With a similar procedure, we derive the gradient associated with $\vec{B}_s(\vec{r})$ (only the x component is shown here)

$$\begin{aligned}\frac{\delta E_{\rm tot}}{\delta B_{s,1}(\vec{r})} \approx {} & \sum_K \sum_{j=1}^{3} \frac{1}{\lambda} \left\{ b_{K,1} \left[A_{s,j} \to A_{s,j} + \lambda \frac{\delta E_{\rm tot}}{\delta a_{K,j}} \right](\vec{r}) - b_{K,1}(\vec{r}) \right\} \\ & + \sum_K \sum_{j=1}^{3} \frac{1}{\lambda} \left\{ b_{K,1} \left[B_{s,j} \to B_{s,j} + \lambda \frac{\delta E_{\rm tot}}{\delta b_{K,j}} \right](\vec{r}) - b_{K,1}(\vec{r}) \right\} \\ & + \sum_K \frac{1}{\lambda} \left\{ b_{K,1} \left[v_s \to v_s + \lambda \frac{\delta E_{\rm tot}}{\delta c_K} \right](\vec{r}) - b_{K,1}(\vec{r}) \right\}. \end{aligned} \tag{17}$$

The gradient associated with $v_s(\vec{r})$ is

$$\begin{aligned}\frac{\delta E_{\rm tot}}{\delta v_s(\vec{r})} \approx {} & \sum_K \sum_{j=1}^{3} \frac{1}{\lambda} \left\{ c_K \left[A_{s,j} \to A_{s,j} + \lambda \frac{\delta E_{\rm tot}}{\delta a_{K,j}} \right](\vec{r}) - c_K(\vec{r}) \right\} \\ & + \sum_K \sum_{j=1}^{3} \frac{1}{\lambda} \left\{ c_K \left[B_{s,j} \to B_{s,j} + \lambda \frac{\delta E_{\rm tot}}{\delta b_{K,j}} \right](\vec{r}) - c_K(\vec{r}) \right\} \\ & + \sum_K \frac{1}{\lambda} \left\{ c_K \left[v_s \to v_s + \lambda \frac{\delta E_{\rm tot}}{\delta c_K} \right](\vec{r}) - c_K(\vec{r}) \right\}. \end{aligned} \tag{18}$$

These gradients [(13), (17), and (18)] can be evaluated efficiently with methods such as spectrum folding.[22] With similar procedures as in the collinear case, the functional derivatives of the KS kinetic energy with respect to $\rho_{\alpha\beta,K}(\vec{r})$ and $\vec{j}_{p,K}(\vec{r})$ are

$$\left.\frac{\delta T_S}{\delta \rho_{\alpha\beta,K}(\vec{r})}\right|_{\{\vec{j}_{p,K}\}} = \varepsilon_K \delta_{\alpha\beta} - v_s(\vec{r})\delta_{\alpha\beta} - \mu_B(\vec{\sigma} \cdot \vec{B}_s(\vec{r}))_{\beta\alpha},$$

$$\left.\frac{\delta T_S}{\delta \vec{j}_{p,K}(\vec{r})}\right|_{\{\rho_{\alpha\beta,K}\}} = -\frac{1}{c}\vec{A}_s(\vec{r}).$$

It is then straightforward to minimize the total energy $E_{\rm tot}$ with respect to KS effective fields.

VI. CONCLUSIONS

In conclusion, we present an efficient direct minimization method for performing OEP calculations. We completely avoid using the time-consuming perturbation theory in the original Yang and Wu formalism[15] and reformulate the problem in terms of easy-to-calculate finite differences. We demonstrate the fast convergence and high accuracy of our method with the non-spin-polarized KS-DFT-EXX calculations on two sodium clusters. Like all other OEP techniques, our approach will encounter numerical difficulties if the KS orbitals do not evolve smoothly with the change of KS effective potential; however, we expect such cases to be exceptional. Our method can be applied immediately to OEP correlation functionals by accounting for both occupied and virtual orbitals, and a straightforward extension to CSDFT has been derived. This efficient direct minimization method should allow OEP to be used more broadly in future for tackling many practical problems where standard exchange-correlation pure density functionals are insufficient.

ACKNOWLEDGMENTS

We acknowledge Stephan Kümmel for providing the Na_8 cluster geometry and thank Florian Libisch for a critical reading of the manuscript. We are grateful for support from the Office of Naval Research and the National Science Foundation.

---

[*] eac@princeton.edu

[1] W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).
[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[3] John P. Perdew and Karla Schmidt, in *Density Functional Theory and Its Applications to Materials*, edited by V. Van Doren, C. Van Alsenoy, and P. Geerlings, AIP Conference Proceedings, Vol. 577 (American Institute of Physics, New York, 2001), pp. 1–20.
[4] S. Kümmel and L. Kronik, Rev. Modern Phys. **80**, 3 (2008).
[5] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).
[6] T. Körzdörfer, M. Mundt, and S. Kümmel, Phys. Rev. Lett. **100**, 133004 (2008).
[7] S. Sharma, J. K. Dewhurst, C. Ambrosch-Draxl, S. Kurth, N. Helbig, S. Pittalis, S. Shallcross, L. Nordström, and E. K. U. Gross, Phys. Rev. Lett. **98**, 196405 (2007).
[8] R. T. Sharp and G. K. Horton, Phys. Rev. **90**, 317 (1953); J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976); A. Görling and M. Levy, *ibid.* **50**, 196 (1994); L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).
[9] S. Hirata, S. Ivanov, I. Grabowski, R. J. Bartlett, K. Burke, and J. D. Talman, J. Chem. Phys. **115**, 1635 (2001).

[10] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **46**, 5453 (1992).

[11] S. Kümmel and J. P. Perdew, Phys. Rev. Lett. **90**, 043004 (2003).

[12] J. H. Rose and H. B. Shore, Solid State Commun. **17**, 327 (1975).

[13] A. Bulgac, C. Lewenkopf, and V. Mickrjukov, Phys. Rev. B **52**, 16476 (1995).

[14] L. Fritsche and J. Yuan, Phys. Rev. A **57**, 3425 (1998).

[15] W. Yang and Q. Wu, Phys. Rev. Lett. **89**, 143002 (2002).

[16] R. A. Hyman, M. D. Stiles, and A. Zangwill, Phys. Rev. B **62**, 15521 (2000).

[17] X. Gonze, Phys. Rev. B **54**, 4383 (1996).

[18] H. Hellmann, *Einführung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), Sec. 54, p. 285; R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[19] X. Gonze, B. Amadon, P. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, P. Ghosez, M. Giantomassi, S. Goedecker, D. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. Verstraete, G. Zerah, and J. Zwanziger, Comput. Phys. Commun. **180**, 2582 (2009).

[20] L. Genovese, T. Deutsch, and S. Goedecker, J. Chem. Phys. **127**, 054704 (2007); L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Beylkin, *ibid.* **125**, 074105 (2006).

[21] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, SIAM, J. Sci. Comput. **16**, 1190 (1995).

[22] L. W. Wang and A. Zunger, J. Chem. Phys. **100**, 2394 (1994).

[23] M. Levy and J. P. Perdew, Phys. Rev. A **32**, 2010 (1985).

[24] F. D. Sala and Andreas Görling, J. Chem. Phys. **116**, 5374 (2002).

[25] U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972); G. Vignale and M. Rasolt, Phys. Rev. Lett. **59**, 2360 (1987); S. Pittalis, Ph.D. thesis, Freie Universität Berlin, 2008.