Conjugate-gradient optimization method for orbital-free density functional calculations

Hong Jiang and Weitao Yang

Conjugate-gradient optimization method for orbital-free density functional calculations

Hong Jiang
*Department of Chemistry, Duke University, Durham, North Carolina 27708-0354
and Department of Physics, Duke University, Durham, North Carolina 27708-0305*

Weitao Yang[a)]
Department of Chemistry, Duke University, Durham, North Carolina 27708-0354

Orbital-free density functional theory as an extension of traditional Thomas-Fermi theory has attracted a lot of interest in the past decade because of developments in both more accurate kinetic energy functionals and highly efficient numerical methodology. In this paper, we developed a conjugate-gradient method for the numerical solution of spin-dependent extended Thomas-Fermi equation by incorporating techniques previously used in Kohn-Sham calculations. The key ingredient of the method is an approximate line-search scheme and a collective treatment of two spin densities in the case of spin-dependent extended Thomas-Fermi problem. Test calculations for a quartic two-dimensional quantum dot system and a three-dimensional sodium cluster Na_{216} with a local pseudopotential demonstrate that the method is accurate and efficient. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1768163]

I. INTRODUCTION

The most attractive feature of density functional theory (DFT) is its use of electron density as the basic variable,[1–3] which is established by the Hohenberg-Kohn theorems.[4] The premodern orbital-free DFT formalism, mainly the Thomas–Fermi–Dirac–von Weizsäcker model,[5–8] though impressive for its simplicity compared to much more complicated wave-function-based approaches, is not quite accurate. Indeed, the great success of modern DFT is due to the introduction of single-particle orbitals in the Kohn-Sham (KS) scheme,[9] which has become the mainstay of applications of the DFT formalism. The pursuit of more accurate orbital-free density functional theory is nevertheless persistent,[4,10–22] and has obtained new momentum in recent efforts in *ab initio* molecular dynamics simulations of complex systems because of its intrinsic linear scaling behavior.[23–29] In this paper, we will use the general term, extended Thomas-Fermi (ETF), to denote all these approaches.

Besides its promising power as a numerical modeling tool, Thomas-Fermi and its extensions also provide the starting point for a lot of other theoretical models.[30–33] For instance, in the Strutinsky model[34] of quantum dots,[33,35] Thomas-Fermi theory accounts for classical charging effects, and quantum effects are incorporated by considering the single-particle interference in the Thomas-Fermi effective potential and the residual interaction between oscillating part of electron density.[33]

In all these cases, an efficient numerical method to solve the ETF equation is a necessary ingredient. The ETF equation can be treated either as a nonlinear self-consistent problem or as a constrained minimization problem. In the former case, a naive implementation of self-consistent iteration is not numerically stable because of the so-called charge-sloshing effect.[36] Elaborate mixing techniques have to be employed to achieve rapid convergence.[11,37] In the latter case, the simplest approach is the steepest-descent (SD) method,[20,36,38] which, though simple to implement, is a well-known "poor" minimization algorithm.[38] Based on analogy to the Kohn-Sham problem, Wang *et al.* formulated the energy minimization in terms of a damped second-order equation.[20] The conjugate-gradient (CG) method is one of the most efficient algorithms for numerical optimization, but its implementation to constrained minimization problems is usually very complicated.[20] Inspired by a direct minimization conjugated-gradient method developed for the Kohn-Sham problem,[36,39,40] in this paper, we develop a CG method to solve the ETF equation, which is simple to implement and very efficient.

The paper is organized as follows: In the following section, the ETF equation is first cast into an illuminating form, followed by a detailed description of the CG method. Section III reports some numerical tests in a two-dimensional quantum dot system and a three-dimensional sodium cluster. The final section summarizes the paper.

II. THEORY

A. Formulation of the problem

According to spin-density functional theory,[1,2] the ground state total energy of a N-electron interacting system in a local external potential $V_{ext}(\mathbf{r})$ is a unique functional of spin densities $\rho_\sigma(\mathbf{r})$ under the constraints,

$$\int \rho_\sigma(\mathbf{r})d\mathbf{r} = N_\sigma, \tag{1}$$

[a)]Author to whom correspondence should be addressed. Electronic mail: weitao.yang@duke.edu

where $\sigma = \alpha, \beta$ denotes spin-up and -down indices, respectively, and N_σ is the number of electrons of each spin, with $N_\alpha + N_\beta = N$. The total energy can be written as (atomic units are used throughout the paper)[4,9]

$$E_{\text{tot}}[\rho_\alpha, \rho_\beta] = T_s[\rho_\alpha, \rho_\beta] + \int V_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}$$
$$+ \frac{1}{2} \int \int \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{\text{xc}}[\rho_\alpha, \rho_\beta], \quad (2)$$

where $T_s[\rho_\alpha, \rho_\beta]$ is the spin-dependent kinetic energy functional of the fictitious noninteracting system that has the same ground state spin densities as the interacting one, and $E_{\text{xc}}[\rho_\alpha, \rho_\beta]$ is the exchange-correlation energy functionals. $T_s[\rho_\alpha, \rho_\beta]$ is related to the spin-conpensated kinetic energy functional, $T_s^0[\rho]$, by[1,41]

$$T_s[\rho_\alpha, \rho_\beta] = \tfrac{1}{2} T_s^0[2\rho_\alpha] + \tfrac{1}{2} T_s^0[2\rho_\beta]. \quad (3)$$

In the standard Kohn-Sham scheme,[1,2,9] T_s^0 is calculated by Kohn-Sham single-particle orbitals, which are eigenfunctions of an effective single-particle Hamiltonian. In the orbital-free ETF theory, however, T_s^0 is approximated as an explicit functional of the density

$$T_s^0[\rho] = T_{\text{TF}}^0[\rho] + \lambda T_{\text{W}}^0[\rho] + T_{\text{nl}}^0[\rho]. \quad (4)$$

The first term is the Thomas-Fermi kinetic energy functional, a local functional of density, $T_{\text{TF}}^0[\rho] = \int t_{\text{TF}}(\rho[\mathbf{r}]) d\mathbf{r}$, where $t_{\text{TF}}(\rho)$ is the kinetic energy density of a homogeneous electron system with electron density ρ,

$$t_{\text{TF}}(\rho) = \begin{cases} \dfrac{\pi}{2} \rho^2 & \text{(2D)} \\[2mm] \dfrac{3}{10} (3\pi^2)^{2/3} \rho^{5/3} & \text{(3D)}. \end{cases} \quad (5)$$

$T_{\text{W}}^0[\rho]$ is the von Weizsäcker functional, which is the exact kinetic energy functional in the limit of rapidly varying densities. In both two-dimensional (2D) and 3D, $T_{\text{W}}^0[\rho]$ takes the following form[1,2,30]

$$T_{\text{W}}^0[\rho] = \frac{1}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r}. \quad (6)$$

λ is an empirical parameter that is widely used to correct the overestimation of the von Weizsäcker term.[1] In this paper, we take $\lambda = 0.25$. The third term in Eq. (4) represents all other nonlocal extensions that have been the interest of recent efforts.[12–20] Since this paper addresses mainly numerical aspects, we will focus on the so-called TF-λW method, which neglects the third term in Eq. (4). The techniques developed here are nevertheless general to other more elaborated kinetic energy functional forms.

Instead of minimizing the total energy directly over the spin densities, we introduce a quantity,[20,23,25,42] $\psi_\sigma(\mathbf{r})$, defined by

$$\rho_\sigma(\mathbf{r}) = [\psi_\sigma(\mathbf{r})]^2, \quad (7)$$

which can be regarded as quasiorbitals. The total energy can then be written as the functional of $\psi_\sigma(\mathbf{r})$,

$$\tilde{E}[\psi_\alpha, \psi_\beta] \equiv E_{\text{tot}}\{[\psi_\alpha(\mathbf{r})]^2, [\psi_\beta(\mathbf{r})]^2\} \quad (8)$$

and the constraint Eq. (1) is transformed to the following normalization condition:

$$\int \psi_\sigma(\mathbf{r})^2 d\mathbf{r} = N_\sigma. \quad (9)$$

Taking the contraint Eq. (9) into account by Lagrange's multiplier method, we define

$$W[\psi_\alpha, \psi_\beta] \equiv \tilde{E}[\psi_\alpha, \psi_\beta]$$
$$- \sum_{\sigma = \alpha, \beta} \tilde{\mu}_\sigma \left\{ \int [\psi_\sigma(\mathbf{r})]^2 d\mathbf{r} - N_\sigma \right\}, \quad (10)$$

where $\tilde{\mu}_\sigma$ is the chemical potential for spin σ electrons.

The advantage of using ψ_σ instead of spin densities as the basic variable is twofold.[20,23,25,42] (1) The requirement that $\rho_\sigma(\mathbf{r})$ has to be positive can be cumbersome to impose in numerical implementations, which is, however, trivial in the case of the ψ_σ formulation; (2) more importantly, by introducing ψ_σ we can transform the ETF problem to exactly the same form as the KS problem so that all efficient techniques that have been developed in past decades for the KS equations can now be easily extended to the ETF problem. The gradient of the total energy with respect to ψ_σ is

$$\frac{\delta W[\psi_\alpha, \psi_\beta]}{\delta \psi_\sigma(\mathbf{r})} = \frac{\delta E[\rho_\alpha, \rho_\beta]}{\delta \rho_\sigma(\mathbf{r})} \frac{\partial \rho_\sigma(\mathbf{r})}{\partial \psi_\sigma(\mathbf{r})} - 2\tilde{\mu}_\sigma \psi_\sigma(\mathbf{r})$$
$$= 2 \left\{ \frac{\delta E[\rho_\alpha, \rho_\beta]}{\delta \rho_\sigma(\mathbf{r})} - \tilde{\mu}_\sigma \right\} \psi_\sigma(\mathbf{r})$$
$$= 2 \left\{ \frac{\delta T_{\text{TF}}}{\delta \rho_\sigma(\mathbf{r})} + \lambda \frac{\delta T_{\text{W}}}{\delta \rho_\sigma(\mathbf{r})} + V_{\text{eff}}^\sigma(\mathbf{r}) - \tilde{\mu}_\sigma \right\} \psi_\sigma(\mathbf{r}),$$
$$(11)$$

where

$$V_{\text{eff}}^\sigma(\mathbf{r}) \equiv V_{\text{ext}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho_\alpha, \rho_\beta]}{\delta \rho_\sigma(\mathbf{r})}. \quad (12)$$

For T_{W} with the form of Eq. (6), we have

$$\frac{\delta T_{\text{W}}}{\delta \rho_\sigma(\mathbf{r})} \psi_\sigma(\mathbf{r}) = - \tfrac{1}{2} \nabla^2 \psi_\sigma(\mathbf{r}), \quad (13)$$

from which, Eq. (11) can be written concisely as

$$\frac{\delta W}{\delta \psi_\sigma(\mathbf{r})} = 2\lambda \{ H_\sigma \psi_\sigma(\mathbf{r}) - \mu_\sigma \psi_\sigma(\mathbf{r}) \}, \quad (14)$$

where $\mu_\sigma \equiv \tilde{\mu}_\sigma / \lambda$, and H_σ is defined as

$$H_\sigma(\mathbf{r}; [\rho_\alpha, \rho_\beta]) = - \frac{1}{2} \nabla^2 + \lambda^{-1} \left\{ \frac{\delta T_{\text{TF}}}{\delta \rho_\sigma(\mathbf{r})} + V_{\text{eff}}^\sigma(\mathbf{r}) \right\}. \quad (15)$$

The variational principle requires $\delta W / \delta \psi_\sigma(\mathbf{r}) = 0$, which leads to

$$H_\sigma \psi_\sigma(\mathbf{r}) = \mu_\sigma \psi_\sigma(\mathbf{r}), \quad (16)$$

which has the same form as the KS equations, but much simplified since there is only one "orbital" for each spin state. As in the KS problem, Eq. (16) is a nonlinear equation that has to be solved in a self-consistent way.

B. A conjugate-gradient method for minimization of ETF total energy

The formulation of the ETF problem using quasiorbitals opens up a lot of new possibilities to solve the ETF problem. As in the case of the KS problem, mainly there are two types of approaches, self-consistent and direct minimization methods.[36,37,40] Here we introduced a direct minimization conjugate-gradient (DMCG) method. Such a method for the KS problem is well developed in plane wave *ab initio* modeling of semiconductor material systems.[36,39] It was further improved by Jiang *et al.* in their KS-DFT study of quantum dots.[40,43]

Starting from an initial guess, a conjugate-gradient algorithm for a numerical optimization problem usually involves three steps:[36,38,40] (1) Calculate the steepest descent vector; (2) construct the conjugate-gradient vector; and finally (3) update the optimization variables by moving along the conjugate-vector direction for a certain distance that is determined either by an exact line search or by approximations. The complication in the ETF problem is the normalization constraint [Eq. (1) or (9)]. The algorithm described below is very similar to what we developed for the KS problem, but there are also some subtle differences that are critical for optimal efficiency. We will describe the algorithm for the spin-dependent case, and its application to the spin-independent case is straightforward.

At the mth iteration, the SD vector is calculated as the negative gradient vector [Eq. (14)] (Dirac's notation for state vectors is used as in Ref. 40)

$$|\zeta_\sigma^{(m)}\rangle = 2\lambda\{\mu_\sigma^{(m)} - H_\sigma^{(m)}\}|\psi_\sigma^{(m)}\rangle, \tag{17}$$

with $\mu_\sigma^{(m)} \equiv \langle\psi_\sigma^{(m)}|H_\sigma^{(m)}|\psi_\sigma^{(m)}\rangle/N_\sigma$. The CG vector is then calculated as

$$|\varphi_\sigma^{(m)}\rangle = |\zeta_\sigma^{(m)}\rangle + \gamma_\sigma^{(m)}|\varphi_\sigma^{(m-1)}\rangle, \tag{18}$$

with

$$\gamma_\sigma^{(m)} = \frac{\langle\zeta_\sigma^{(m)}|\zeta_\sigma^{(m)}\rangle}{\langle\zeta_\sigma^{(m-1)}|\zeta_\sigma^{(m-1)}\rangle}, \tag{19}$$

for $m>1$ and $\gamma_\sigma^{(m)}=0$ for $m=1$. To satisfy the normalization constraint of ψ_σ, the CG vector is further orthogonalized to $|\psi_\sigma^{(m)}\rangle$ and normalized to N_σ,

$$|\varphi_\sigma'^{(m)}\rangle = \left(1 - \frac{|\psi_\sigma^{(m)}\rangle\langle\psi_\sigma^{(m)}|}{N_\sigma}\right)|\varphi_\sigma^{(m)}\rangle, \tag{20}$$

$$|\phi_\sigma^{(m)}\rangle = |\varphi_\sigma'^{(m)}\rangle\left(\frac{N_\sigma}{\langle\varphi'^{(m)}_\sigma|\varphi'^{(m)}_\sigma\rangle}\right)^{1/2}. \tag{21}$$

ψ_σ is then updated by

$$|\psi_\sigma^{(m+1)}\rangle = |\psi_\sigma^{(m)}\rangle\cos\theta_\sigma^{\min} + |\phi_\sigma^{(m)}\rangle\sin\theta_\sigma^{\min}, \tag{22}$$

where the values of $\theta_\sigma^{\min}$ are determined by minimizing the total energy as a function of θ_α and θ_β,

$$E(\theta_\alpha,\theta_\beta) \equiv E_{\text{tot}}[\rho_\alpha(\mathbf{r};\theta_\alpha),\rho_\beta(\mathbf{r};\theta_\beta)]$$

$$\equiv \tilde{E}[\psi_\alpha(\mathbf{r};\theta_\alpha),\psi_\beta(\mathbf{r};\theta_\beta)], \tag{23}$$

with

$$\psi_\sigma(\mathbf{r};\theta_\sigma) = \psi_\sigma^{(m)}(\mathbf{r})\cos\theta_\sigma + \phi_\sigma^{(m)}(\mathbf{r})\sin\theta_\sigma \tag{24}$$

and $\rho_\sigma(\mathbf{r};\theta_\sigma) = [\psi_\sigma(\mathbf{r};\theta_\sigma)]^2$, which is equivalent to minimizing the Lagrangian $W[\psi_\alpha,\psi_\beta]$ since the normalization constraints are intrinsically imposed here.

The first and second derivatives of $E(\theta_\alpha,\theta_\beta)$ with respect to θ_σ can be obtained by

$$\frac{\partial E(\theta_\alpha,\theta_\beta)}{\partial\theta_\sigma} = \int d\mathbf{r}\,\frac{\delta\tilde{E}}{\delta\psi_\sigma(\mathbf{r};\theta_\sigma)}\frac{\partial\psi_\sigma(\mathbf{r};\theta_\sigma)}{\partial\theta_\sigma}$$

$$= 2\lambda\langle\dot\psi_\sigma|H_\sigma(\theta_\alpha,\theta_\beta)|\psi_\sigma(\theta_\sigma)\rangle \tag{25}$$

and

$$\frac{\partial^2 E(\theta_\alpha,\theta_\beta)}{\partial\theta_\sigma\partial\theta_{\sigma'}} = 2\lambda\left\{\langle\ddot\psi_\sigma(\theta_\sigma)|H_\sigma(\theta_\alpha,\theta_\beta)|\psi_\sigma(\theta_\sigma)\rangle\delta_{\sigma,\sigma'}\right.$$

$$+\langle\dot\psi_\sigma(\theta_\sigma)|H_\sigma(\theta_\alpha,\theta_\beta)|\dot\psi_\sigma(\theta_\sigma)\rangle\delta_{\sigma,\sigma'}$$

$$\left.+\langle\dot\psi_\sigma(\theta_\sigma)|\frac{\partial H_\sigma(\theta_\alpha,\theta_\beta)}{\partial\theta_{\sigma'}}|\psi_\sigma(\theta_\sigma)\rangle\right\}, \tag{26}$$

with

$$H_\sigma(\theta_\alpha,\theta_\beta) \equiv H(\mathbf{r};[\rho_\alpha(\theta_\alpha),\rho_\beta(\theta_\beta)]), \tag{27}$$

$$\frac{\partial H_\sigma(\theta_\alpha,\theta_\beta)}{\partial\theta_{\sigma'}} = \lambda^{-1}\left\{\left(2t''_{\text{TF}}[2\rho_\sigma(\theta_\sigma)]\delta_{\sigma,\sigma'}\right.\right.$$

$$\left.+\frac{\delta^2 E_{\text{XC}}[\rho_\alpha,\rho_\beta]}{\delta\rho_\sigma\delta\rho_{\sigma'}}\right)\dot\rho_{\sigma'}(\theta'_\sigma)$$

$$\left.+\int d\mathbf{r}'\frac{\dot\rho_{\sigma'}(\mathbf{r}';\theta_\sigma)}{|\mathbf{r}-\mathbf{r}'|}\right\}, \tag{28}$$

$$\dot\psi_\sigma(\theta_\sigma) \equiv \frac{\partial\psi_\sigma(\theta_\sigma)}{\partial\theta_\sigma} = -\psi_\sigma^{(m)}\sin\theta_\sigma + \phi_\sigma^{(m)}\cos\theta_\sigma, \tag{29}$$

$$\ddot\psi_\sigma(\theta_\sigma) \equiv \frac{\partial^2\psi_\sigma(\theta_\sigma)}{\partial\theta_\sigma^2} = -\psi_\sigma(\theta_\sigma), \tag{30}$$

$$\dot\rho_\sigma(\theta_\sigma) \equiv \frac{\partial\rho_\sigma(\theta_\sigma)}{\partial\theta_\sigma} = 2\psi_\sigma(\theta_\sigma)\dot\psi_\sigma(\theta_\sigma). \tag{31}$$

There are a lot of standard optimization techniques available to solve this two-variable minimization problem (single-variable minimization for the spin-independent ETF).[38,44] We will, however, pursue an approximate scheme similar to the method we used in the KS case.[40] We first transform Eq. (25) into a more illuminating form

$$\frac{\partial E(\theta_\alpha,\theta_\beta)}{\partial\theta_\sigma} = 2\lambda\langle-\psi_\sigma^{(m)}\sin\theta_\sigma + \phi_\sigma^{(m)}\cos\theta_\sigma|$$

$$\times H_\sigma(\theta_\alpha,\theta_\beta)|\psi_\sigma^{(m)}\cos\theta_\sigma + \phi_\sigma^{(m)}\sin\theta_\sigma\rangle$$

$$= \lambda[-A_\sigma(\theta_\alpha,\theta_\beta)\sin 2\theta_\sigma$$

$$+ B_\sigma(\theta_\alpha,\theta_\beta)\cos 2\theta_\sigma] \tag{32}$$

with

$$A_\sigma(\theta_\alpha,\theta_\beta) \equiv \langle \psi_\sigma^{(m)} | H_\sigma(\theta_\alpha,\theta_\beta) | \psi_\sigma^{(m)} \rangle$$
$$- \langle \phi_\sigma^{(m)} | H_\sigma(\theta_\alpha,\theta_\beta) | \phi_\sigma^{(m)} \rangle \qquad (33)$$

and

$$B_\sigma(\theta_\alpha,\theta_\beta) \equiv 2 \langle \phi_\sigma^{(m)} | H_\sigma(\theta_\alpha,\theta_\beta) | \psi_\sigma^{(m)} \rangle, \qquad (34)$$

where we have used the fact that both ψ_σ and ϕ_σ are real. Now we introduce the approximation

$$H_\sigma(\theta_\alpha,\theta_\beta) \simeq H_\sigma(0,0), \qquad (35)$$

so that by setting the first derivate, Eq. (32), to be zero, we obtain

$$\theta_\sigma^{\min} \simeq \tfrac{1}{2} \tan^{-1} \frac{B_\sigma(0,0)}{A_\sigma(0,0)}. \qquad (36)$$

The validity of Eq. (35) can be established by the following observations: The dominant parts of the effective Hamiltonian H_σ are the kinetic energy operator and external potential, which are independent of θ_σ [See Eq. (15)]; the other parts of H_σ depend on θ_σ through $\rho_\sigma(\theta_\sigma)$, which plays a weaker role in determining $\theta_\sigma^{\min}$. In the Appendix, we formulate a more accurate approximation that goes one step beyond Eq. (35), in which the dependence of the Hartree potential on θ_σ is incorporated.

The algorithm described above is very close to the band-by-band DMCG method used in Kohn-Sham calculations.[36,39,40] There is, however, one critical difference. In the spirit of the band-by-band scheme, one might treat two spin densities alternatively: Iterate only one spin density at a time for N_{band} times with the other spin density fixed. Such a sequential treatment has the disadvantage that the conjugate-gradient relaxation is disrupted every N_{band} times, which impairs partly the high-efficiency intrinsic to a conjugate-gradient algorithm. We also found in our numerical tests that the approximate line-minimization scheme using Eq. (36) is sometimes numerically unstable, and a more accurate approximation to $\theta^{\min}$ like that described in the Appendix or an "exact" line search is required. In spite of that, the sequential scheme is still much more efficient than the steepest descent method; when an exact line search is required, Eq. (36) provides a very accurate initial guess. We will denote that treatment as the SCG (for sequential conjugate gradient) method in the following section. In contrast, the method described above iterates the two spin densities simultaneously, and in doing so has taken full advantage of the fact that the two quasispin orbitals are not required to be orthogonal to each other. In this treatment, the approximation made in Eq. (36) is found to be stable and accurate, which dramatically reduces the computation efforts; iterations are always carried out in the conjugate-gradient direction. Because the normalization constraint is imposed in each iteration, there is no accumulation of numerical errors that may occur in some CG algorithms. We will denote this treatment as CCG (for concurrent conjugate-gradient). For the spin-independent case, these two approaches are identical.
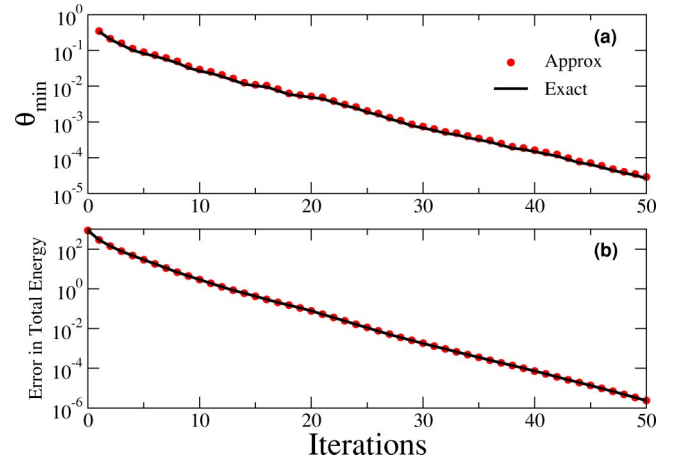

FIG. 1. Comparison of approximate and exact line search methods in spin-independent ETF. (a) $\theta_{\min}$ calculated by the approximate method, Eq. (36) (dot), and by exact Brent's method in a typical ETF calculation (line); (b) convergence errors vs iteration number using approximate and exact line search methods.

III. NUMERICAL TESTS

We will report numerical results only for finite systems, but conclusions from these test calculations should be applicable to periodic infinite systems as well. For other important components of a full implementation of the ETF theory to real systems, we essentially use same techniques as we did in the KS case:[40] We use the particle-in-the-box basis for the representation of quasiorbitals, ψ_σ, which is a variant of plane-wave basis for finite systems; the action of the effective Hamiltonian on quasiorbitals, $H_\sigma \psi_\sigma$ is effected by fast-sine transform;[40] for the Hartree potential, we use the Fourier convolution method on a doubly extended grid.[40,45] Local spin-density approximation is used for E_{xc}. In particular, we use the Tanatar-Ceperley functional[46] for 2D systems and the Vosko-Wilk-Nusair functional[47] for 3D systems.

We first test the performance of the method in a 2D quantum dot model system with a coupled quartic oscillator potential (QOP),[43,48]

$$V_{\text{ext}}(x,y) = a\left(\frac{x^4}{b} + by^4 - 2\lambda x^2 y^2 + \gamma(x^2 y - xy^2)r \right), \qquad (37)$$

with $r = \sqrt{x^2+y^2}$, $a = 10^{-4}$, $b = \pi/4$, $\lambda = 0.53$, and $\gamma = 0.2$. This potential was used by the authors in numerical studies of e-e interaction effects in quantum dots.[40,43] The numerical results are obtained for $N = 200$. For the spin-dependent case, we consider the triplet state, i.e., $N_\alpha = 101$ and $N_\beta = 99$.

We first test the accuracy of the approximation of Eq. (36) by comparing with the exact line search in the spin-independent ETF case. The exact $\theta^{\min}$ is calculated by the standard Brent's method,[38] in which Eq. (36) is used to calculate the initial guess. Figure 1(a) shows approximate and exact $\theta^{\min}$ in a typical minimization process. The approximate $\theta^{\min}$ remains in good agreement with the exact one through the iterations. Figure 1(b) plots the convergence er-
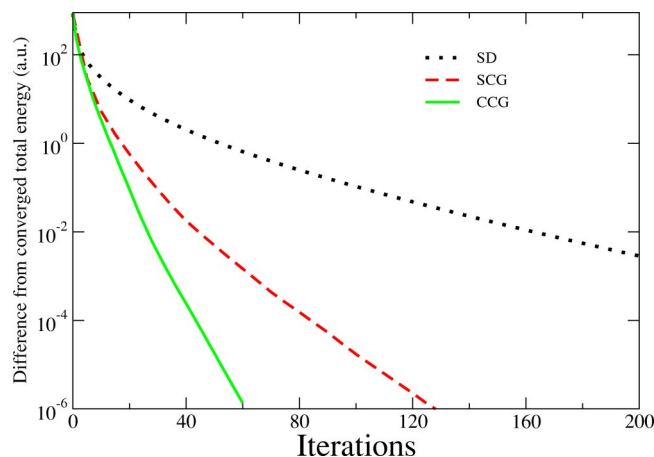
FIG. 2. (Color online) Error in the total energy vs iteration number in three different minimization methods for spin-dependent ETF: steepest descent (dot), sequential conjugate gradient (dash), and collective conjugate gradient (solid). In SCG, Brent's method for line minimization is used, and N_{band} $=5$.

rors versus iteration number using approximate and exact line minimization methods, respectively. The agreement between them is almost perfect.

Figure 2 shows the convergence behaviors of three different minimization methods, SD, SCG, and CCG, in the case of the spin-dependent ETF calculations. For the SD method, we use the formalism in Ref. 20, but with a exact line search. Using the approximation in Eq. (36), the computation cost for a single step in CCG is much less than that of SD and SCG, yet the convergence is still faster to attain in CCG than in SCG as well as in SD. The combination of these two improving factors reduces the computation efforts in CCG by almost two order of magnitude with respect to SCG.
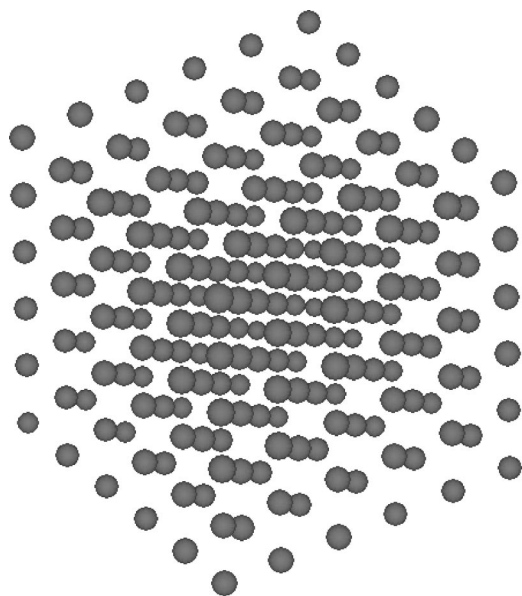

FIG. 3. The geometry of the Na cluster used in the test calculations. Na atoms are located in a finite simple cubic lattice. The distance between neighboring atoms are taken as 4 a.u.
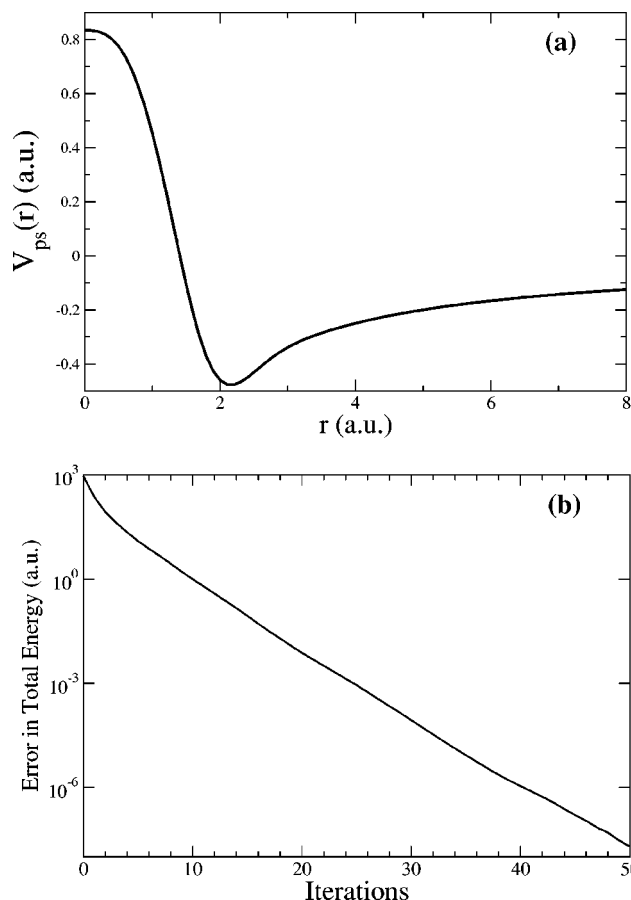

FIG. 4. (a) Local ion pseudopotential for Na used in test calculations of Na_{216}; (b) error in the total energy in the Na_{216} system vs iteration number.

Finally, to test the performance of our method in more realistic systems, we consider a sodium cluster, Na_{216}. Since our main purpose here is to test the performance of our method, we take the geometry of the cluster as a cubic without relaxation, as illustrated in Fig. 3; the distance between neighboring Na atoms is taken as 4 a.u. In this case, the external potential is formed by the superposition of the local pseudopotential for the Na ion on different sites,

$$V_{ext}(\mathbf{r}) = \sum_i V_{ps}(|\mathbf{r}-\mathbf{R}_i|). \tag{38}$$

For $V_{ps}(r)$ we use Bachelet-Hamann-Schlüter's norm-conserving pseudopotential[49] neglecting the nonlocal part, which is illustrated in Fig. 4(a). In the test calculations, we use a cubic box in real space with the length of 28.0 a.u. and 81 mesh points in each direction. We take both the change in the total energy between successive iterations, $\Delta E \equiv E^{(m)} - E^{(m-1)}$, and the norm of the gradient as the convergence criteria.

Figure 4(b) shows the convergence behavior of our method in this realistic system. As in the previous 2D model system, the convergence can be easily attained in about 50 iterations, which confirms the high efficiency of our conjugate-gradient method in realistic molecular systems.

IV. SUMMARY AND DISCUSSION

In this paper, an efficient conjugate-gradient method for direct minimization of the extended Thomas-Fermi total energy is proposed. The method is inspired by a similar approach for the Kohn-Sham problem. The key ingredient of the method is an approximate line-search scheme and a collective treatment of two spin densities in the case of spin-dependent ETF problem. The high performance of the method was verified in a simple 2D model system and a realistic sodium cluster.

We close the paper with two comments. First, though the method presented in the paper is based on transforming the ETF problem into a KS-type form, which is enabled by the presence of the von Weizsäcker term, our method has more general significance in terms of its mathematical structure. We note that the validity of the formulation in Sec. II B does not depend on the exact form of the effective Hamiltonian H_σ. By defining

$$\lambda H_\sigma \psi_\sigma \equiv \frac{1}{2} \frac{\delta \tilde{E}}{\delta \psi_\sigma}, \tag{39}$$

our method can be easily generalized to other orbital-free DFT formalisms with or without the von Weizsäcker term.

Second, recently in a benchmark ETF studies on atomic and diatomic systems using Gaussian basis sets, Chan *et al.* found that all gradient-based methods including CG and quasi-Newton methods perform poorly in minimizing ETF total energy.[21] Though we tested our method only in the plane-wave type representation, the formulation of the method is nevertheless general, and therefore should be equally applicable to local basis represented systems. We will leave further investigations to future studies.

ACKNOWLEDGMENTS

APPENDIX: AN IMPROVED LINE MINIMIZATION SCHEME

The approximation in Eq. (36) can be improved by considering the dependence of the Hartree potential on θ. To simplify the notation, we present the formulation only for the spin-independent case, and the spin index is therefore dropped. Instead of neglecting the dependence of H on θ completely [Eq. (35)], we make a less dramatic approximation,

$$H(\mathbf{r};\theta) \simeq H(\mathbf{r};0) + \int \frac{\Delta\rho(\mathbf{r}';\theta)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \tag{A1}$$

where $\Delta\rho(\mathbf{r};\theta)$ is the change of the density

$$\Delta\rho(\mathbf{r};\theta) \equiv \rho(\mathbf{r};\theta) - \rho(\mathbf{r};0)$$
$$= [\psi^{(m)}(\mathbf{r})\cos\theta + (\mathbf{r})\sin\theta]^2 - [\psi^{(m)}(\mathbf{r})]^2$$
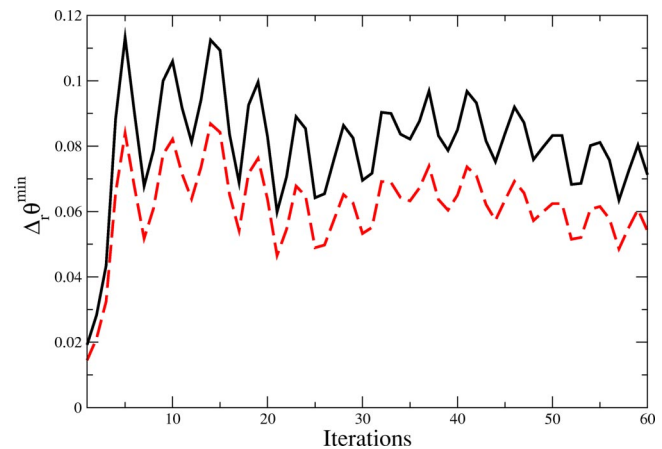$$= -\chi_1(\mathbf{r})\sin^2\theta + \chi_2(\mathbf{r})\cos\theta\sin\theta \tag{A2}$$

with

FIG. 5. Relative errors in $\theta^{\min}$ [Eq. (A11) using Eq. (36) (solid) and Eq. (A10) (dashed), respectively] in the spin-independent ETF calculation of the QOP system with $N=200$. The exact $\theta^{\min}$ is calculated by the standard Brent's method (Ref. 38).

$$\chi_1(\mathbf{r}) \equiv [\psi^{(m)}(\mathbf{r})]^2 - [\phi^{(m)}(\mathbf{r})]^2, \tag{A3}$$

$$\chi_2(\mathbf{r}) \equiv 2\psi^{(m)}(\mathbf{r})\phi^{(m)}(\mathbf{r}). \tag{A4}$$

Using Eqs. (A1) and (A2), we can obtain, after some straightforward transformation,

$$A(\theta) = A_0 - C_{11}\sin^2\theta + C_{12}\sin\theta\cos\theta, \tag{A5}$$

$$B(\theta) = B_0 - C_{12}\sin^2\theta + C_{22}\sin\theta\cos\theta, \tag{A6}$$

with

$$A_0 \equiv \langle\psi^{(m)}|H(0)|\psi^{(m)}\rangle - \langle\phi^{(m)}|H(0)|\phi^{(m)}\rangle, \tag{A7}$$

$$B_0 \equiv 2\langle\phi^{(m)}|H(0)|\psi^{(m)}\rangle, \tag{A8}$$

$$C_{ij} \equiv \int\int \frac{\chi_i(\mathbf{r})\chi_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \quad \text{for } i,j=1,2. \tag{A9}$$

$\theta^{\min}$ is then determined by requiring

$$\frac{\delta E(\theta)}{\delta\theta} \simeq \lambda\left\{ -\left(A_0 - \frac{C_{11}}{2}\right)\sin 2\theta + \left(\frac{B_0 - C_{12}}{2}\right)\cos 2\theta \right. $$
$$\left. - \frac{C_{11}-C_{22}}{4}\sin 4\theta + \frac{C_{12}}{2}\cos 4\theta \right\}, \tag{A10}$$

to be zero. Though there is no simple close expression for the root of Eq. (A10) such as Eq. (36), it can be easily solved numerically using standard techniques such as the Newton-Raphson method.[38] The increase in the computation efforts using this more accurate line-minimization approximation is marginal when compared to that using Eq. (36).

We tested the line minimization method in the QOP system with $N=200$ in the spin-independent case. Figure (5) shows the relative errors in $\theta^{\min}$

$$\Delta_r\theta^{\min} \equiv \frac{\theta^{\min}_{\text{approx}} - \theta^{\min}_{\text{exact}}}{\theta^{\min}_{\text{exact}}} \tag{A11}$$

during a typical calculation, which shows that the line-minimization method based on Eq. (A10) significantly improved the accuracy of the approximated $\theta^{\min}$. On the other hand, the number of iterations required to achieve conver-

gence remains the same, which confirms further the validity of Eq. (36). We note, however, that the improvement of Eq. (A10) with respect to Eq. (36) is important in the SCG scheme for spin-dependent ETF calculations to guarantee the numerical stability, as discussed in the paper. We expect that the improved method will also be useful in systems where Eq. (36) might fail.

[1] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[2] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer, Berlin, 1990).

[3] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[4] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

[5] L. H. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927).

[6] E. Fermi, Acad. Naz. Lincei **6**, 602 (1927).

[7] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).

[8] L. H. Weizsäcker, Proc. Cambridge Philos. Soc. **23**, 542 (1935).

[9] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

[10] J. A. Alonso and L. A. Girifalco, Phys. Rev. B **17**, 3735 (1978).

[11] W. Yang, Phys. Rev. A **34**, 4575 (1986).

[12] A. E. DePristo and J. D. Kress, Phys. Rev. A **35**, 438 (1987).

[13] H. Lee, C. Lee, and P. R. G., Phys. Rev. A **44**, 768 (1991).

[14] B. Wang, M. J. Stott, and U. von Barth, Phys. Rev. A **63**, 052501 (2001).

[15] L. W. Wang and M. P. Teter, Phys. Rev. B **45**, 13196 (1992).

[16] M. Brack and R. K. Bhaduri, *Semiclassical Physics* (Addison-Wesley, Reading, MA, 1997).

[17] P. Garcia-Gonzalez, J. E. Alvarellos, and E. Chacon, Phys. Rev. A **54**, 1897 (1996).

[18] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **58**, 13465 (1998).

[19] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **60**, 16350 (1999).

[20] Y. A. Wang and E. A. Carter, in *Theoretical Methods in Condensed Phase Chemistry*, edited by S. D. Schwartz (Kluwer, Dordrecht, 2000), pp. 117–184.

[21] G. K. Chan, A. J. Cohen, and N. C. Handy, J. Chem. Phys. **114**, 631 (2001).

[22] F. Perrot, J. Phys.: Condens. Matter **6**, 431 (1994).

[23] V. Shah, D. Nehete, and D. G. Kanhere, J. Phys.: Condens. Matter **6**, 10773 (1994).

[24] E. Smargiassi and P. A. Madden, Phys. Rev. B **49**, 5220 (1994).

[25] N. Govind, J. Wang, and H. Guo, Phys. Rev. B **50**, 11175 (1994).

[26] P. Blaise, S. A. Blundell, and C. Guet, Phys. Rev. B **55**, 15856 (1997).

[27] M. Foley and P. A. Madden, Phys. Rev. B **53**, 10589 (1996).

[28] A. Domps, P. G. Reinhard, and E. Suraud, Phys. Rev. Lett. **80**, 5520 (1998).

[29] S. C. Watson and E. A. Carter, Comput. Phys. Commun. **128**, 67 (2000).

[30] Y. Wang, J. Wang, H. Guo, and E. Zaremba, Phys. Rev. B **52**, 2738 (1995).

[31] S. Sinha, R. Shankar, and M. V. N. Murthy, Phys. Rev. B **62**, 10896 (2000).

[32] J. Shi and X. C. Xie, Phys. Rev. Lett. **88**, 086401 (2002).

[33] D. Ullmo and H. U. Baranger, Phys. Rev. B **64**, 245324 (2001).

[34] V. M. Strutinsky, Nucl. Phys. A **122**, 1 (1968).

[35] D. Ullmo, T. Nagano, S. Tomosovic, and H. U. Baranger, Phys. Rev. B **63**, 125339 (2001).

[36] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992), and references therein.

[37] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996), and references therein.

[38] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterlin, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1989).

[39] M. P. Teter, M. C. Payne, and D. C. Allan, Phys. Rev. B **40**, 12255 (1989).

[40] H. Jiang, H. U. Baranger, and W. Yang, Phys. Rev. B **68**, 165337 (2003).

[41] G. L. Oliver and J. P. Perdew, Phys. Rev. A **20**, 397 (1979).

[42] N. Govind, Ph.D. thesis, McGill University, 1995.

[43] H. Jiang, H. U. Baranger, and W. Yang, Phys. Rev. Lett. **90**, 026806 (2003).

[44] D. P. Bertsekas, *Nonlinear Programming* (Athena, Belmont, MA, 1999).

[45] G. J. Martyna and M. E. Tuckerman, J. Chem. Phys. **110**, 2810 (1999).

[46] B. Tanatar and D. M. Ceperley, Phys. Rev. B **39**, 5005 (1989).

[47] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 100 (1980).

[48] O. Bohigas, S. Tomsovic, and D. Ullmo, Phys. Rep. **223**, 43 (1993).

[49] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).