

Ab initio linear scaling method for electronic structure calculations via local scaling transformations

D. G. Kanhere* and Ajeeta Dhavale†

Department of Physics, University of Pune, Pune 411 007, India

E. V. Ludeña and V. Karasiev

Chemistry Center, Venezuelan Institute for Scientific Research, IVIC, Apartado 21827, Caracas 1020-A, Venezuela

(Received 22 May 2000; published 10 November 2000)

An algorithm based on local scaling transformations for electronic structure calculations that scales linearly with the size of the system is presented. The key feature of the method is the absence of the orthogonalization step during iterative minimization. We illustrate the feasibility and potential of the method by applying it to total energy calculations for a variety of small clusters, viz., Na₂, Na₇Al, Na₂₀, Si₄, and Al₁₃. The method is easily parallelizable and therefore has the potential to deal with large real life systems.

PACS number(s): 36.40.Cg, 71.24.+q, 31.15.-p, 71.15.-m

It is now well recognized that microscopic many-body quantum mechanical calculations yielding detailed information about electronic structure are at the heart of understanding the properties of systems such as molecules, clusters, surfaces, and solids. Typically such calculations employ either Hartree-Fock and related techniques or, for larger systems, the powerful density functional theory (DFT) [1]. Most of the popular algorithms use either direct diagonalization or iterative minimization schemes. In either of these cases the algorithm scales as the cube of the system size; the cost comes from the Householder rotations or from the orthogonalization step. It can be seen that even with the availability of faster computers saturation will be reached very quickly. Hence, there is considerable interest in developing linear scaling [$O(N)$] methods, i.e., methods scaling as $T=C_1N$, where T is the computer time and N the size of the system. Indeed, such $O(N)$ methods are considered absolutely essential for understanding complex real life materials involving very large numbers of atoms.

In the last few years there has been some progress in this direction. A number of methods exhibiting such linear scaling have been proposed [2]. However, almost all successful implementations rely upon the localized nature of orbitals. It has been pointed out that such methods implemented with localization constraints, e.g., the orbital minimization method, lead to unacceptably large minimization iterations due to multiple minima. These methods are dependent on the fact that the density matrix is diagonally dominant and the off-diagonal elements decay with distance, the decay rate being system dependent. This is the nearsightedness principle of Kohn [3]. Most of these methods employ tight binding schemes although a few applications have been made using the density functional method on real life systems. A detailed review of $O(N)$ methods has been presented by Goedecker [4].

Since in the conventional methods the limiting N^3 behavior arises from the orthogonalization step, it is desirable to

seek a method of energy minimization that avoids this step. In this Brief Report we present one such algorithm by taking a completely different path from those mentioned above. Our algorithm is based on local scaling density functional theory (LSDFT) as advanced by Ludeña and co-workers [5]. The LSDFT is a rigorous formulation of DFT, constructive in nature, and satisfies the N and ν representability conditions on the energy functional. In principle, the method is applicable to Hartree-Fock or Kohn-Sham Hamiltonians and yields the corresponding orbitals and energies. In what follows we give a short presentation of the formulation relevant to our algorithm. The formal aspects, generalizations, and other details of the method may be found elsewhere [5].

Local scaling transformations (LST's) are coordinate transformations that carry a vector $\vec{r} \rightarrow \vec{r}\lambda(\vec{r})$, where $\lambda(\vec{r})$ is a scalar function. The Jacobian for this transformation is given by

$$J(\vec{r}; \vec{r}\lambda(\vec{r})) = \lambda^3(\vec{r}) \{1 + \vec{r} \cdot \vec{\nabla}_{\vec{r}} \ln[\lambda(\vec{r})]\}. \quad (1)$$

Let us consider an orthonormal set of orbitals $\{\phi_g(\vec{r})\}$ called the generating set and the associated generating density

$$\rho_g(\vec{r}) = \sum_i |\phi_g^i(\vec{r})|^2. \quad (2)$$

Let $\rho(\vec{r})$ be the density homotopic to $\rho_g(\vec{r})$; then we can construct a local scaling transformation such that

$$\rho(\vec{r}) = J(\vec{r}; \vec{r}\lambda(\vec{r})) \rho_g(\vec{r}\lambda(\vec{r})). \quad (3)$$

It is easy to see that the transformed orbitals obtained from

$$\psi_i(\vec{r}) = [J(\vec{r}; \vec{r}\lambda(\vec{r}))]^{1/2} \phi_g^i(\vec{r}\lambda(\vec{r})) \quad (4)$$

generate the transformed density $\rho(\vec{r})$ via

$$\rho(\vec{r}) = \sum_i |\psi_i(\vec{r})|^2. \quad (5)$$

*Electronic address: kanhere@unipune.ernet.in

†Electronic address: ajd@physics.unipune.ernet.in

The Jacobian must be positive definite for the transformation to be unique. The most crucial property of this transformation relevant to our discussion is that it maintains the orthonormality of the orbitals. Such a local scaling transformation can be viewed as a coordinate transformation that continuously deforms $\rho_g(\vec{r})$ to obtain $\rho(\vec{r})$. The scalar function $\lambda(\vec{r})$ connecting $\rho_g(\vec{r})$ and $\rho(\vec{r})$ satisfies the nonlinear equation

$$\lambda^3(\vec{r}) = \frac{\rho(\vec{r})}{\{1 + \vec{r} \cdot \vec{\nabla} \ln[\lambda(\vec{r})]\} \rho_g(\vec{r} \lambda(\vec{r}))}. \quad (6)$$

Now consider the Hohenberg-Kohn total energy functional for a N_e -electron system in an external potential $V(\vec{r})$. The total energy for this system is given by

$$E[\{\psi_i(\vec{r})\}] = T[\{\psi_i(\vec{r})\}] + E_{Coul}[\rho] + E_{xc}[\rho] + E_{ext}[\rho], \quad (7)$$

where E_{Coul} is the electron-electron Coulomb interaction energy, E_{xc} is the exchange correlation energy, and E_{ext} is the electron-ion interaction energy.

In all the conventional methods the total energy E is minimized with respect to $\psi_i(\vec{r})$ by explicitly incorporating the constraint of orthonormality of orbitals. Now we note that $\{\psi_i(\vec{r})\}$ is related to the generating set $\{\phi_g^i(\vec{r})\}$ via the local scaling transformation as

$$\psi_i(\vec{r}) = (\lambda^3(\vec{r}) \{1 + \vec{r} \cdot \vec{\nabla} \ln[\lambda(\vec{r})]\})^{1/2} \phi_g^i(\vec{r} \lambda(\vec{r})). \quad (8)$$

Hence the total energy E can be viewed as a functional of $\phi_g(\vec{r})$ and $\lambda(\vec{r})$. This immediately suggests a minimization scheme for the total energy by considering it as a functional of $\lambda(\vec{r})$ with some suitable fixed generating set $\{\phi_g(\vec{r})\}$, i.e., $E[\lambda(\vec{r}); \{\phi_g(\vec{r})\}]$. The ground-state energy E_0 is then attained as the extremum of a double minimization procedure,

$$E_0 = \min_{\lambda(\vec{r})} \min_{\{\phi_g(\vec{r})\}} E[\lambda(\vec{r}); \{\phi_g(\vec{r})\}]. \quad (9)$$

Interorbit minimization is minimization at fixed $\lambda(\vec{r})$:

$$E_0 \leq \min_{\{\phi_g(\vec{r})\}} E[\lambda(\vec{r}); \{\phi_g(\vec{r})\}] \quad (10)$$

and intraorbit minimization is minimization at fixed $\{\phi_g(\vec{r})\}$:

$$E_0 \leq \min_{\lambda(\vec{r})} E[\lambda(\vec{r}); \{\phi_g(\vec{r})\}]. \quad (11)$$

A steplike procedure combining these two minimization schemes will eventually lead to E_0 . The minimization proposed here is the intraorbit minimization [Eq. (11)], carried out with the constraint that the Jacobian of the transformation is positive definite. In fact our results show that, with the proper choice of generating orbitals, the intraorbit minimization procedure is sufficient to attain the ground-state energy, at least to within acceptable accuracy. In spite of the formulation being available for quite some time, there are no reported applications except for simple atoms to our knowledge. This is mainly because the transformation was viewed

as one converting the trial charge density to the known exact ground-state density of the system. This leads to the problem of determination of $\lambda(\vec{r})$ from a nonlinear equation [Eq. (6)]. Note that the nonlinear equation has to be solved with the constraint on the Jacobian, which could be very tricky.

It is now possible to set up an efficient algorithm by drawing upon the considerable expertise developed in iterative minimization schemes such as conjugate gradient methods. Here we present a simple steepest descent based algorithm.

LSDFT algorithm. Initialization: Choose an initial orthonormal set $\{\phi_i(\vec{r})\}$, $\lambda(\vec{r}) = 1$ and $\psi_i(\vec{r}) = \phi_i(\vec{r})$. Iterate the following steps till minimization. (1) Calculate $\rho(\vec{r})$, E and $\delta E / \delta \lambda(\vec{r})|_{\lambda(\vec{r})=1}$. (2) Calculate the new $\lambda(\vec{r}) = 1 - \alpha \delta E / \delta \lambda(\vec{r})|_{\lambda(\vec{r})=1}$. (3) Calculate the new $\psi_i(\vec{r}) = [J(\vec{r}; \vec{r} \lambda(\vec{r}))]^{1/2} \phi_i(\vec{r} \lambda(\vec{r}))$. (4) Set $\phi_i(\vec{r}) = \psi_i(\vec{r})$.

In the above algorithm α is evaluated by a standard line minimization technique, i.e., $E(\alpha)$ is line minimized to obtain the best α and the resulting $\lambda(\vec{r})$ is accepted only if the corresponding Jacobian is positive definite. Otherwise a smaller value of α yielding lower energy $E(\alpha)$ is to be accepted. In practice, for all the cases investigated this has never happened.

For implementation of the above algorithm we need to evaluate the functional derivative $\delta E / \delta \lambda(\vec{r})$. It is easy to show that

$$\frac{\delta E}{\delta \lambda(\vec{r})} = \sum_i \frac{\delta E[\{\psi_i\}]}{\delta \psi_i^*(\vec{r})} \frac{\delta \psi_i^*(\vec{r})}{\delta \lambda(\vec{r})}, \quad (12)$$

$$\left. \frac{\delta E}{\delta \lambda(\vec{r})} \right|_{\lambda(\vec{r})=1} = \sum_i (H - \epsilon_i) \psi_i(\vec{r}) \left[\frac{3}{2} \phi_i^*(\vec{r}) + \vec{r} \cdot \vec{\nabla} \phi_i^*(\vec{r}) \right], \quad (13)$$

where H is the standard Kohn-Sham Hamiltonian and $\epsilon_i = \langle \psi_i | H | \psi_i \rangle$. The second term in the square brackets was obtained by Taylor series expansion assuming that the change in $\lambda(\vec{r})$ from its initial value is small. In fact, this is the most convenient way to ensure the positivity constraint on J .

We have successfully tested this algorithm by calculating the ground-state energies of a number of clusters. The implementation is based on the standard plane wave pseudopotential formulation using a periodically repeated large unit cell. It may be mentioned that the algorithm is easy to incorporate within existing pseudopotential codes. The numerical details and other aspects of the algorithm will be reported elsewhere [6]. We report the total energies for the clusters Na_2 , Na_7Al , Na_{20} , Si_4 , and Al_{13} , representing 2, 10, 20, 16, and 39 electron systems. The timings reported here are those for total energies obtained to within 0.005 hartree of the Kohn-Sham total energies with identical parameters such as energy cutoff, cell dimension, pseudopotential, etc. All the calculations except for Si use the local Bachelet-Hamann-Schluter pseudopotential [7].

The CPU timings in units of timings for Na_2 as a function of the number of orbitals for all the clusters studied are

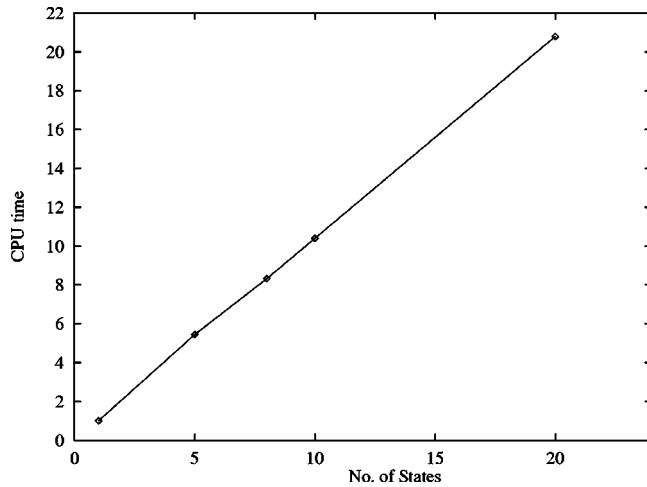


FIG. 1. The CPU timings in units of timings for Na_2 as a function of the number of orbitals for all the clusters.

shown in Fig. 1. The timings required to generate the initial guess are not included. As expected the graph shows a linear behavior.

Figure 2 shows the generating density $\rho_g(\vec{r})$ (shown by a dashed line) and the final density $\rho(\vec{r})$ (shown by a continuous line). Figures 3(a), 3(b), and 3(c) show the generating wave functions for the first state, eighth state, and highest occupied state (all shown by a dashed line), and the corresponding final wave functions (shown by a continuous line). All the plots are along the x axis for the Na_{20} cluster. These illustrate the ability of LST to continuously deform the initial density and wave functions toward the final self-consistent density and wave functions in about 40–50 iterations. The final density and wave functions are virtually indistinguishable from the full self-consistent Kohn-Sham density.

Now we offer some remarks about the practical aspects of the algorithm.

(1) The procedure is sensitive to the choice of generating orbitals. These must be chosen to be homotopic, i.e., to have

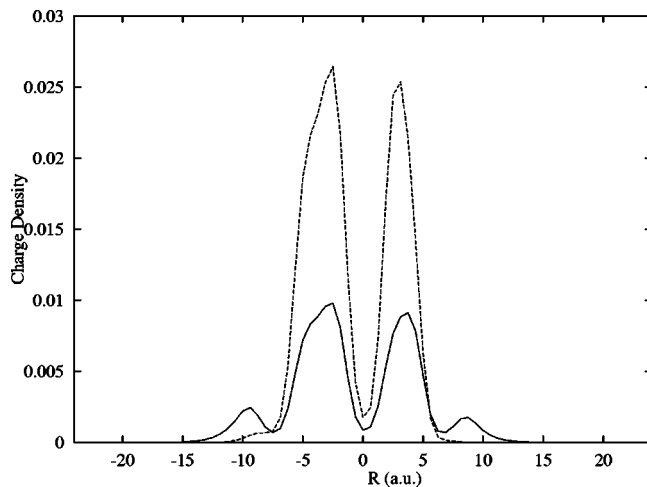


FIG. 2. The generating charge density (dashed line) and the final charge density (continuous line) for Na_{20} cluster along the x axis.

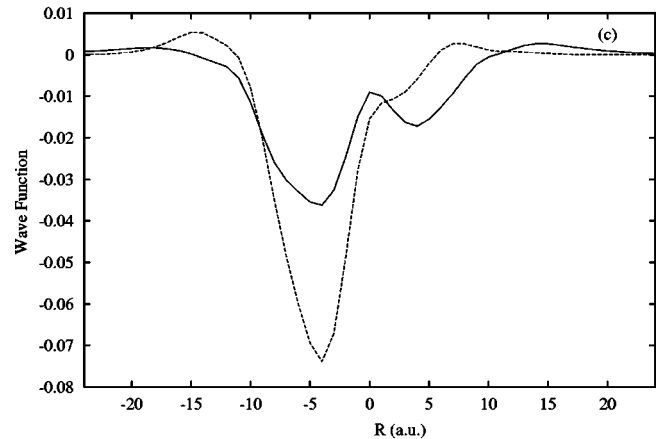
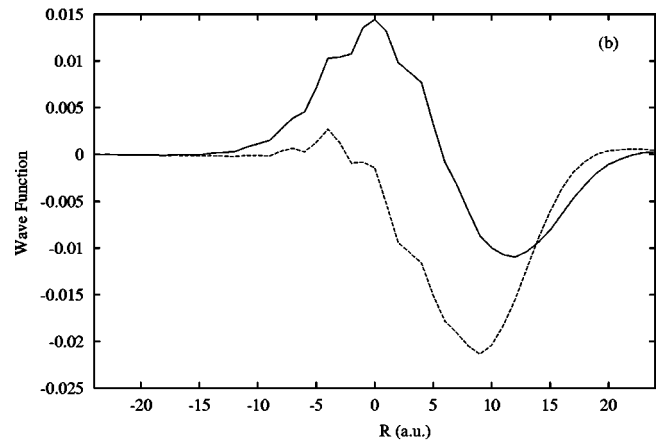
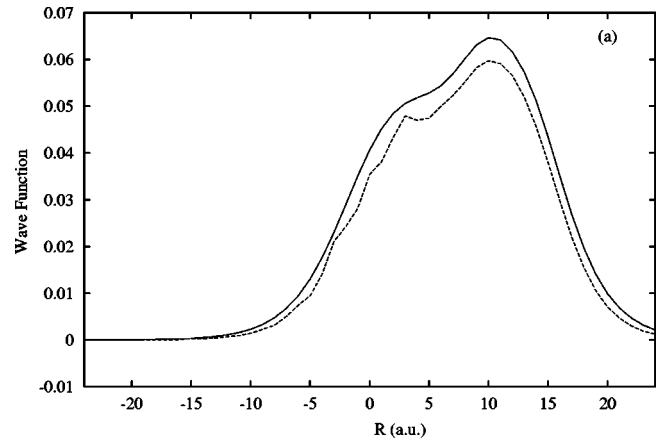


FIG. 3. The generating wave function (dashed line) and the final wave function (continuous line) corresponding to the first state (a), the eighth state (b), and the highest occupied state (c) for the Na_{20} cluster along the x axis.

the same number of minima/maxima and the same nodal structure as the final orbitals. Fortunately, for small clusters this is not a serious problem since a number of methods are available for generating a good initial guess. In the present work we have used the orbitals obtained after a few (typically two or three) iterations of the standard Kohn-Sham scheme starting with a random set of orbitals. However, for large systems careful attention will have to be given to the proper choice of the generating set of orbitals.

(2) Although in principle LST preserves orthonormality, in practice this is true only approximately because of discretization of the numerical mesh and the interpolation procedure required in Eq. (4). In most cases we have seen that $\langle \psi_i | \psi_j \rangle$ is of the order of 10^{-4} and improves as the minimization proceeds. However, it is desirable to reorthogonalize the current set of orbitals after some fixed number of iterations. Fortunately, it is required to be done only a few times (two or three) during the entire minimization process.

(3) In all the cases considered we have achieved the accuracy mentioned above in about 40–50 iterations.

(4) The convergence of the method slows down considerably after about 0.005 Hartree. A possible reason is that by choosing fixed initial orbitals and varying $\lambda(\vec{r})$ we are carrying out only an intraorbit minimization and are neglecting the interorbit step. Of course, there are also numerical errors due to the interpolation required [Eq. (4)] and the use of the steepest descent method. However, an additional few conventional Kohn-Sham iterations (which combine both inter- and intraorbital minimization) have always allowed us to achieve an accuracy better than 0.001 Hartree.

(5) In some cases, albeit rarely, it is observed that the convergence stagnates after a few (~ 10) iterations. This can be remedied either by improving the initial guess or by jumping to another orbit by performing one or two Kohn-

Sham iterations. The use of reorthogonalization and a small number of standard Kohn-Sham iterations adds N^2 scaling although with a very small prefactor.

(6) An attractive feature of this algorithm is that it lends itself to parallelization in a very natural way. Once $\lambda(\vec{r})$ is determined, all orbitals are scaled independently and their contribution to the total energy can be calculated independently.

Finally, we note that the calculation of forces required for the dynamics is identical to plane wave methods and scales as N^2 .

In conclusion, we have presented an algorithm with nearly linear scaling $O(N)$ that is based on local scaling transformations. In principle, the method scales linearly because it avoids orthogonalization of orbitals. The method has been successfully applied to obtain the ground-state total energies of several clusters and the results bring out the potential of this method for large scale systems.

It is a pleasure to acknowledge a number of helpful discussions with P. V. Panat, R. K. Pathak, and R. R. Zope. Financial assistance from the Department of Science and Technology is gratefully acknowledged. One of us (A.D.) acknowledges CSIR, India. E.V.L. acknowledges the support of Conicet of Venezuela through G-97000741.

-
- [1] M. Payne, M. Teter, D. Allan, T. Aris, and J. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).
 [2] U. Stephan and D. Drabold, *Phys. Rev. B* **57**, 6391 (1998); X.P. Li, W. Nunes, and D. Vanderbilt, *ibid.* **47**, 10 891 (1993); F. Mauri, G. Galli, and R. Car, *ibid.* **47**, 9973 (1993); P. Ordejón, D.A. Drabold, R.M. Martin, and M.P. Grumbach, *ibid.* **51**, 1456 (1995); F. Mauri and G. Galli, *ibid.* **50**, 4316 (1994); S. Goedecker and L. Colombo, *Phys. Rev. Lett.* **73**, 122 (1994).
 [3] W. Kohn, *Phys. Rev. Lett.* **76**, 3168 (1996).
 [4] S. Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999).
 [5] E.S. Kryachko and E.V. Ludeña, *Phys. Rev. A* **43**, 2179

- (1991); R. López-Boada, E.V. Ludeña, V. Karasiev, and R. Pino, *J. Chem. Phys.* **107**, 6722 (1997); T. Koga, Yoshiaki Yamamoto, and E. Ludeña, *Phys. Rev. A* **43**, 5814 (1991); E. Ludeña and R. López-Boada, *Top. Curr. Chem.* **180**, 169 (1996).
 [6] Ajeeta Dhavale and D.G. Kanhere (unpublished).
 [7] G.B. Bachelet, D.R. Hamann, and M. Schluter, *Phys. Rev. B* **26**, 4199 (1982). For Si we have used a semiempirical local pseudopotential given by J.A. Appelbaum and D.R. Hamann [*Phys. Rev. B* **8**, 1777 (1973)]. It may be noted that the inclusion of nonlocality is straightforward and does not pose any problems.