PHYSICAL REVIEW B, VOLUME 64, 195126

Convergent recursive O(N) method for *ab initio* tight-binding calculations

T. Ozaki[1,2] and K. Terakura[1,3]
[1]*RICS, National Institute of Advanced Industrial Science and Technology (AIST), central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*
[2]*JRCAT-ATP, central 4, 1-1-1 Higashi, Tsukuba, Ibaraki 305-0046, Japan*
[3]*JRCAT-AIST, central 4, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8562, Japan*

(Received 31 July 2001; published 25 October 2001)

A theory is presented for a recursion method for O(N) *ab initio* tight-binding calculations based on the density-functional theory. A long-standing problem of generalizing the recursion method to a non-orthogonal basis, which is a crucial step to make the recursion method applicable to *ab initio* calculations, is solved by the introduction of hybrid representation and a two-sided block Lanczos algorithm. As a test of efficiency of the proposed method, convergence properties in energy and force of insulators, semiconductors, metals, and molecules are studied for not only simple model systems but also some real materials within the density-functional theory. The present O(N) method possesses good convergence properties for metals as well as insulators.

DOI: 10.1103/PhysRevB.64.195126 PACS number(s): 71.15.Mb, 71.15.Nc, 71.15.Pd

I. INTRODUCTION

The application of the conventional *ab initio* electronic structure calculations to large systems is hampered by their inherent O(N^3) scaling properties with N the number of atoms. In order to overcome this difficulty, efficient linear-scaling algorithms, which are referred to as O(N) methods, have been developed during the last decade.[1–3] A few applications of these O(N) methods have actually demonstrated the power of these O(N) methods.[4] However, several problems still remain in these O(N) methods. In particular, it is well known that the variational O(N) methods produce large errors in the energy of metals,[5] since in both metals and narrow-gap semiconductors the density matrix ρ has long-range correlations in real space compared to that of insulators with a wide gap.[6] Therefore, the application of the O(N) methods relying on the locality of ρ is restricted to the materials with a wide gap.

Recently, one of the authors has shown that the block bond-order potential (BOP) is a convergent moment-based O(N) method that provides good convergence properties for energy and force in metals as well as insulators and semiconductors within an orthogonal tight-binding (TB) representation.[2] Although it had been well documented that the moment-based methods cannot reproduce the vacancy in diamond or silicon within a low number of moments,[5,7] the block BOP gave the first convergent results for vacancies in insulators using a moment-based method with a low number of moments,[2] which means that this O(N) method is the most accurate approach within the moment-based methods. The block algorithm guarantees that the σ and π bonds are treated properly and that the band energy is invariant for the rotation of systems,[8] and the terminator in the Green functions can estimate the long range contributions in a most effective way[9] for the total energy and force calculations. Thus, the block BOP may be applicable to a wide variety of materials with reliable accuracy. However, in order to make the method applicable to *ab initio* calculations, we have to remove the limitation of orthogonal basis and reformulate the method in nonorthogonal basis. Several attempts have been made to generalize the recursion method to nonorthogonal basis.[9] Nevertheless, within our knowledge of the traditional recursion method the formalism is not in a satisfactory level and only very few examples of *ab initio* calculations are available,[10] while the generalized Fermi-operator-expansion method by Stephan and Drabold[11] is one of the moment-based methods that can be successfully applied to nonorthogonal orbitals.

In this paper we present a generalized block BOP for the nonorthogonal TB basis by introducing a hybrid representation and the two-sided block Lanczos algorithm. We will show convergence properties of the method within density-functional theory (DFT) and demonstrate that the present method is a promising and practical O(N) method in both insulators and metals.

II. THEORY

It will be assumed that one-particle wave functions are expanded with a nonorthogonal localized basis set ($|i\alpha\rangle$) where i is a site index, and α an orbital index. Then, the density of electrons $\rho(\mathbf{r})$ in nonspin-polarized systems can be written as

$$\rho(\mathbf{r})=\sum_{i\alpha,j\beta}\chi_{i\alpha}(\mathbf{r})\chi_{j\beta}(\mathbf{r})\Theta_{i\alpha,j\beta}, \tag{1}$$

where $\chi_{i\alpha}(\mathbf{r})\equiv\langle\mathbf{r}|i\alpha\rangle$, and the bond-order $\Theta_{i\alpha,j\beta}$ is related to the imaginary part of the one-particle Green function as follows:

$$\Theta_{i\alpha,j\beta}=-\frac{2}{\pi}\mathrm{Im}\int G_{i\alpha,j\beta}(E+\mathrm{i}0^{+})f\left(\frac{E-\mu}{k_BT}\right)dE \tag{2}$$

with the Fermi function $f(x)=1/[1+\exp(x)]$, where 0^{+} represents a positive infinitesimal. $G_{i\alpha,j\beta}(Z)$ is an expectation value of Greenian $\hat{G}(Z)\equiv(Z-\hat{H})^{-1}$ for dual bases $|\widetilde{i\alpha}\rangle$ and $|\widetilde{j\beta}\rangle$ defined by

$$|\widetilde{i\alpha}\rangle=\sum_{j\beta}|j\beta\rangle S^{-1}_{j\beta,i\alpha}, \tag{3}$$

where $S^{-1}_{j\beta,i\alpha}$ is the inverse of overlap matrix $S_{i\alpha,j\beta} \equiv\langle i\alpha|j\beta\rangle$. From Eqs. (1) and (2), we see that the functionals such as the total energy in DFT can be reformulated as the functionals of the bond order. Therefore, we concentrate on evaluating the bond order with reasonable accuracy and great reduction in the computational efforts.

Let us introduce a hybrid representation of Hamiltonian that is a non-Hermitian matrix represented by the original and the dual bases as $H'_{i\alpha,j\beta}=\langle\widetilde{i\alpha}|\hat{H}|j\beta\rangle$. The hybrid Hamiltonian can be written in the matrix form as $H'=S^{-1}H$, where $H_{i\alpha,j\beta}\equiv\langle i\alpha|\hat{H}|j\beta\rangle$. With the relation $G(Z)(ZS-H) =\mathrm{I}$, the hybrid Green function $G'(Z)$ defined by

$$G'_{i\alpha,j\beta}(Z)=\{G(Z)S\}_{i\alpha,j\beta}=\langle\widetilde{i\alpha}|\hat{G}(Z)|j\beta\rangle \tag{4}$$

satisfies $G'(Z)(Z\mathrm{I}-H')=\mathrm{I}$. One of the merits of using $G'(Z)$ is that its diagonal element gives directly the Mulliken population $P_{i\alpha}$ of an orbital $|i\alpha\rangle$

$$\begin{aligned}P_{i\alpha}&=-\frac{2}{\pi}\mathrm{Im}\int G'_{i\alpha,i\alpha}(E+0^{+})f\left(\frac{E-\mu}{k_BT}\right)dE\\&=\sum_{j\beta}\Theta_{i\alpha,j\beta}S_{j\beta,i\alpha}.\end{aligned} \tag{5}$$

In the block BOP, determination of the chemical potential is needed to conserve the total number of electrons N_{ele} in the system,[2] so that the relation of Eq. (5) is very advantageous to computational efficiency because of the simple relation $N_{\mathrm{ele}}=\Sigma_{i\alpha}P_{i\alpha}$. Thus, we present below a prescription how to calculate the hybrid Green functions. The diagonal elements of the Green-function matrix can be calculated in a numerically stable way by the recursion method[9] based on the Lanczos algorithm.[12] The block-BOP method is a general recursion method for evaluating efficiently both the diagonal and off-diagonal elements of the Green-function matrix by the recursion method. Moreover the use of a single site containing all the localized orbitals as the starting state in the *block* Lanczos algorithm rather than a single orbital in the usual one conserves the rotational invariance of the total energy. In the present case of nonorthogonal basis, we further extend the formalism to adopt a two-sided block Lanczos algorithm,[13] since the hybrid Hamiltonian is no more Hermitian. The set of central equations is

$$\begin{aligned}&\hat{H}|U_n)=|U_n)\underline{A}_n+|U_{n-1})\underline{B}_n+|U_{n+1})\underline{C}_{n+1},\\&(\tilde{U}_n|\hat{H}=\underline{A}_n(\tilde{U}_n|+\underline{C}_n(\tilde{U}_{n-1}|+\underline{B}_{n+1}(\tilde{U}_{n+1}|.\end{aligned} \tag{6}$$

$\underline{A}_n$, $\underline{B}_n$, and $\underline{C}_n$ are recursion block coefficients with $M_i \times M_i$ in size, where M_i is the number of localized orbitals on the starting atom i, and the underline indicates that the element is a block. In the two-sided block Lanczos algorithm the Lanczos vectors in the left and right sides have a biorthogonality relation. It is essential to start the two-sided block Lanczos algorithm with a single site and its corresponding dual state as

$$\begin{aligned}&|U_0)=(|i1\rangle,|i2\rangle,\ldots,|iM_i\rangle),\\&|\tilde{U}_0)=(|\widetilde{i1}\rangle,|\widetilde{i2}\rangle,\ldots,|\widetilde{iM_i}\rangle).\end{aligned} \tag{7}$$

Equation (7) is an optimum choice in terms of computational accuracy and efficiency because of the rotational invariance of the total energy and the consistent description of the different properties of σ, π, and δ bonds.

In the Lanczos basis representation the Hamiltonian H^L is block tridiagonalized as a non-Hermitian matrix and the Green-function matrix $G^L(Z)$ is the inverse of the matrix $(Z\mathrm{I}-H^L)$, so that the block diagonal element $\underline{G}^L_{00}(Z) =(\tilde{U}_0|\hat{G}|U_0)$ can be written explicitly in the form of the multiple inverse as follows:

$$\underline{G}^L_{00}(Z)=[Z\underline{\mathrm{I}}-\underline{A}_0-\underline{C}_1[Z\underline{\mathrm{I}}-\underline{A}_1-\underline{C}_2[\cdots]^{-1}\underline{B}_2]^{-1}\underline{B}_1]^{-1}, \tag{8}$$

where the index L indicates the representation based on the Lanczos basis. The off-diagonal elements of hybrid Green-function matrix can be calculated by using a recurrence relation that can be derived basically along the same line as that described for the case of orthogonal basis.[2] The explicit expression consistent with Eq. (8) is given below:

$$\begin{aligned}\underline{G}^L_{0n}(Z)=&[\underline{G}^L_{0n-1}(Z)(Z\underline{\mathrm{I}}-\underline{A}_{n-1})-\underline{G}^L_{0n-2}(Z)\underline{C}_{n-1}-\delta_{1n}\underline{\mathrm{I}}]\\&\times(\underline{B}_n)^{-1},\end{aligned} \tag{9}$$

where δ is Kronecker's delta, and $\underline{G}_{0-1}(Z)=\underline{C}_0=\underline{0}$. The block elements of the Green-function matrix have the same relation to the bond orders based on the Lanczos basis $\underline{\Theta}^L_{0n}$ as Eq. (2) of the dual basis representation. Therefore, we can obtain the bond orders of Eq. (2) through the following transformation:

$$\underline{\Theta}_{ij}=\sum_{n,k}\underline{\Theta}^L_{0n}\underline{\tilde{U}}_{nk}\underline{S}^{-1}_{kj}, \tag{10}$$

where $\underline{\tilde{U}}_{nj}$ is defined by $\underline{\tilde{U}}_{nj}=(\tilde{U}_n|(|j1\rangle,|j2\rangle,\ldots,|jM_j\rangle)$. As a result of the simple inverse transformation Eq. (10), we only have to perform the evaluation and the integration of the Green functions of the zeroth block line in the Lanczos basis representation, which means that the computational time of the algorithm is about two times[14] longer compared to that of the orthogonal case.[2] Only the hybrid representation can provide this simple relation Eq. (10) as well as Eq. (5), while the other representations suffer from computational inefficiency.[9,10] In the generalized block BOP using the nonorthogonal basis we need to calculate S^{-1}, the inverse of the overlap matrix. In the following calculations, we used a different O(N) efficient method for inverting the overlap matrix.[15]

III. CONVERGENCE PROPERTIES

In Fig. 1 we show convergence properties of the band energy in an insulator and a metal described by a simple *s*-valent TB as a test of the present method. The errors in the band energy at the seven-shell cluster and recursion levels are 0.2% and 0.9% for the insulator and the metal, respec-

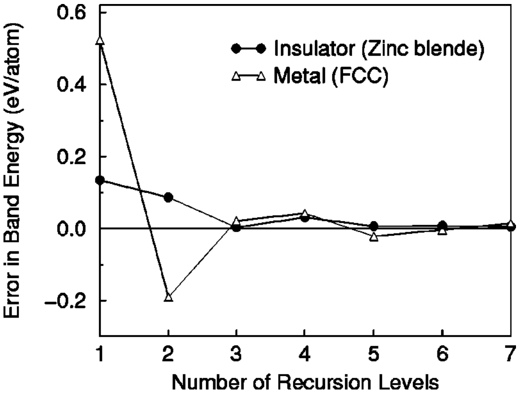

FIG. 1. Error, with respect to the standard *k*-space calculations, in the band energy for an insulator (zinc blende) and a metal (fcc) described by a simple *s*-valent TB model in which the nearest neighbor hopping and overlap integrals are -1.0 eV and 0.1, respectively, with others being zero, and the number of electrons is the same as that of atoms. The zinc blende has a direct gap of 1.0 eV that was controlled by the gap of the on-site energies of the different atoms. In these calculations, the seven-shell cluster and a square-root terminator were used.

tively. Thus, we see that the block BOP gives sufficient convergent results in both the simple insulator and metal. Figures 2(a) and 2(b) show the error in the band energy at the five-shell cluster and recursion levels for insulators and metals described by a simple *s*-valent TB as a function of direct band gap and electronic temperature, respectively. In insulators the error goes to zero as the gap increases, while the errors, whose absolute values are no more than 0.5% compared to the band energy in the whole region, are relatively small. In metals the error becomes almost negligible for the higher electronic temperature. This behavior in both insulators and metals is consistent with the recent study about the locality of the density matrix,[6] though the block BOP depends on the convergence of the moment expansions for the density matrix rather than the locality of the density matrix.[2] From the comparison in the NaCl and fcc structures it is clear that the use of the terminator in the diagonal Green functions effectively reduces the error in both cases.

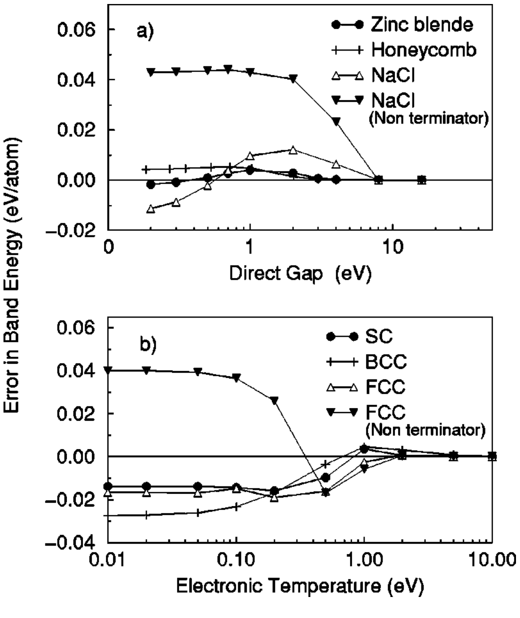

FIG. 2. Error in the band energy for (a) insulators and (b) metals, calculated at the five-shell cluster and recursion levels. The calculations were carried out with the same *s*-valent TB model as that in Fig. 1 using a square-root terminator. For NaCl and fcc the nonterminator results are also shown.

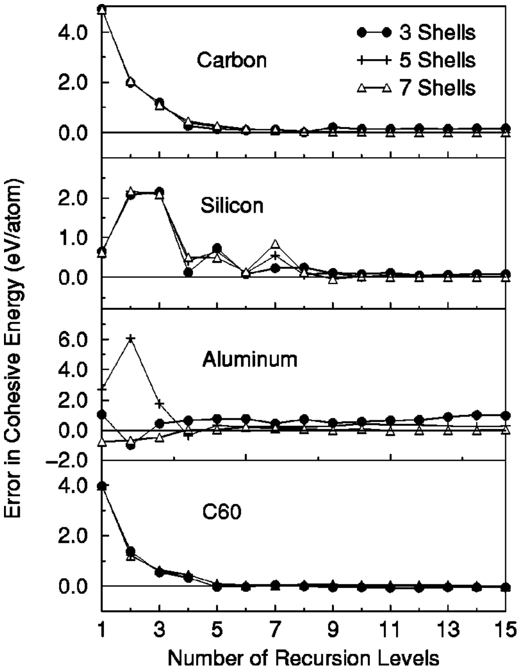

FIG. 3. Error in the cohesive energy for carbon in the diamond structure, silicon in the diamond structure, fcc aluminum, and C_{60} for three-, five-, and seven-shell clusters, calculated using a square-root terminator. These calculations were performed within DFT.

Next, we discuss convergence properties of the block BOP in realistic materials within the TB-based DFT proposed by Sankey and Niklewski.[16] Figure 3 shows the con-

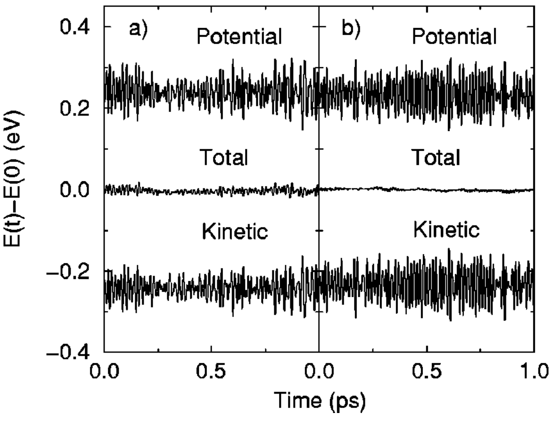

FIG. 4. Potential, kinetic, and total energies as a function of time for molecular dynamics simulations of carbon using a three-shell cluster and a square-root terminator. The fixed unit cell contains eight atoms. In panels (a) and (b) the results are for five and ten recursion levels at 300 K, respectively. The time step is 0.5 fs.

vergence properties of the cohesive energy for carbon in the diamond structure, silicon in the diamond structure, fcc aluminum, and C_{60} molecule.[17] In carbon and silicon the cohesive energies rapidly converge to the *k*-space results in the five- and seven-shell clusters, while the convergence values for the three-shell cluster are in error by 0.4% and 0.9% from the *k*-space results, respectively. Even for metallic aluminum, the convergence is very fast with respect to the number of recursion levels and the errors in the converged values are only 0.3% and 0.1% for the three- and five-shell clusters, respectively. For C_{60} the convergence is achieved with the three-shell cluster. The error at the sixth recursion level is only 0.02%. As a test of the consistency between the total energy and the forces, constant-energy molecular dynamics simulations have been performed for diamond within DFT. In Fig. 4 we show the energy for diamond at 300 K using five and ten recursion levels. For five and ten recursion levels we see that the total energy is almost conserved, while the ten-recursion-level calculation gives a better result. Although the simulations have been done for eight atoms, the small number of atoms is sufficient to demonstrate the consistency between the total energy and forces. These test calculations indicate that the block BOP can give forces consistent with the total energy.

IV. CONCLUSIONS

In summary, we have presented a theory of the block BOP for O(N) *ab initio* TB calculations. The introduction of the hybrid representation and the two-sided block Lanczos algorithm enables us to generalize the theory to a nonorthogonal basis set in a natural way. The test calculations for the simple *s*-valent TB systems and some real systems within DFT suggest that the O(N) method provides rapid convergence properties for metals as well as insulators with sufficient accuracy. Thus, we conclude that the block BOP is a robust O(N) method that is applicable to a wide variety of materials in the *ab initio* TB approach.

ACKNOWLEDGMENTS

We would like to thank Y. Morikawa and H. Kino for helpful suggestions about the DFT calculations.

---

[1] D. G. Pettifor, Phys. Rev. Lett. **63**, 2480 (1989); M. Aoki, *ibid.* **71**, 3842 (1993); A. P. Horsfield, A. M. Bratkovsky, D. G. Pettifor, and M. Aoki, Phys. Rev. B **53**, 1656 (1996); A. P. Horsfield, A. M. Bratkovsky, M. Fearn, D. G. Pettifor, and M. Aoki, *ibid.* **53**, 12 694 (1996).

[2] T. Ozaki, Phys. Rev. B **59**, 16 061 (1999); T. Ozaki, M. Aoki, and D. G. Pettifor, *ibid.* **61**, 7972 (2000).

[3] S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994); W. Yang, *ibid.* **66**, 1438 (1991); G. Galli and M. Parrinello, *ibid.* **69**, 3547 (1992); F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993); X.-P. Li, R. W. Nunes, and D. Vanderbilt, *ibid.* **47**, 10 891 (1993); M. S. Daw, *ibid.* **47**, 10 895 (1993).

[4] G. Galli and F. Mauri, Phys. Rev. Lett. **73**, 3471 (1994); A. Canning, G. Galli, and J. Kim, *ibid.* **78**, 4442 (1997); T. Ozaki, Y. Iwasa, and T. Mitani, *ibid.* **84**, 1712 (2000); S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999), and references therein.

[5] D. R. Bowler, M. Aoki, C. M. Goringe, A. P. Horsfield, and D. G. Pettifor, Modell. Simul. Mater. Sci. Eng. **5**, 199 (1997).

[6] S. Ismail-Beigi and T. Arias, Phys. Rev. Lett. **82**, 2127 (1999).

[7] J. D. Kress and A. F. Voter, Phys. Rev. B **52**, 8766 (1995); A. F. Voter, J. D. Kress, and R. N. Silver, *ibid.* **53**, 12 733 (1996).

[8] R. Jones and M. W. Lewis, Philos. Mag. B **49**, 95 (1984); J. Inoue and Y. Ohta, J. Phys. C **20**, 1947 (1987).

[9] R. Haydock, V. Heine, and M. J. Kelly, J. Phys. C **5**, 2845 (1972); **8**, 2591 (1975); R. Haydock, Solid State Phys. **35**, 216 (1980).

[10] A. Gibson, R. Haydock, and J. P. LaFemina, Phys. Rev. B **47**, 9229 (1993).

[11] U. Stephan and D. Drabold, Phys. Rev. B **57**, 6391 (1998).

[12] C. Lanczos, J. Res. Natl. Bur. Stand. **45**, 225 (1950).

[13] Z. Bai, D. Day, and Q. Ye, SIAM J. Matrix Anal. Appl. **20**, 1060 (1999), and references therein.

[14] For diamond, the computational time on a sharp PC-MJ140M taken per MD step per atom are 1.8 and 2.8 sec for the orthogonal and nonorthogonal cases, respectively.

[15] We have recently developed an efficient inversion method of the overlap matrix in which the computational effort is proportional to the number of atoms. The details of the method are found in cond-mat/0106048 (unpublished).

[16] O. F. Sankey and D. J. Niklewski, Phys. Rev. B **40**, 3979 (1989).

[17] In these density-functional calculations minimal-basis sets for valence electrons were used for carbon, silicon, and aluminum atoms. The radii of the radial-wave function confinement are 2.1, 3.1, and 3.7 Å, respectively. The minimal-basis sets give 1.253 (1.244), 2.280 (2.246), and 2.515 (2.466) Å as an equilibrium bond length of dimer for carbon, silicon, and aluminum, respectively, where the values in the parentheses are experimental results.