PHYSICAL REVIEW B VOLUME 59, NUMBER 24 15 JUNE 1999-II

# Bond-order potential based on the Lanczos basis

Taisuke Ozaki
*Japan Advanced Institute of Science and Technology, Tatsunokuchi, Ishikawa 923-1292, Japan*

(Received 14 December 1998; revised manuscript received 25 January 1999)

A general recursion method for tight-binding molecular dynamics simulations is described in terms of a bond-order potential based on the Lanczos basis. The simple recursive algorithm for calculating the band energy and the forces is intrinsically linear in the scaling of the computational efforts for large systems and very suitable for parallel computation. As a test of this method, constant energy molecular dynamics simulations are performed for carbon materials. The conserved total energy indicates that the forces are of good quality. [S0163-1829(99)06623-0]

The tight-binding (TB) model is practically invaluable for the simultaneous investigation of electronic structure and molecular dynamics in a wide range of systems from covalent materials to transition metals. In order to apply the TB model to real complex systems, for example, disordered materials, liquids, and large molecules, we need to perform simulations on very large systems. However, it exceeds the capacity of modern computers to treat large systems, including thousands of atoms, using widely known methods for solving the electronic structure problem, such as the conjugate-gradient method, because the computational efforts scale as the third power of system size. Therefore, several efficient schemes with linear scaling algorithms have been proposed during the last decade. [1–8] Among them, the bond-order potential (BOP) method [1–4] can give insight into bonding and structure of molecules and solids. Also the method has an advantage in terms of parallel computation in the same way as other $O(N)$ methods such as the density matrix method and the Fermi operator expansion method. [6–8] The BOP method derives the local density of states (LDOS) and the bond orders, which are related to the band energy and the forces, respectively, from the local atomic environment. We can easily parallelize the BOP program code, since the calculation of the band energy and the forces of each atom is highly independent.

However, the BOP method has a complicate algorithm compared with conventional recursion methods, because it is a delicate problem to evaluate off-diagonal elements of the Green's-function matrix which are related to the bond orders. To calculate the bond orders, the method adopts procedures where the off-diagonal elements, obtained by means of auxiliary Hilbert space, are corrected by sum rules following from the identity $(ZI-H)G(Z)=I$.[2–4] Therefore it is obvious that the complexity will place hurdles on the road to applications for large scale molecular dynamics simulations in spite of its role as a tool to analyze the origin of the bonding and structure and the advantage for parallel computation. A new BOP formalism should be derived in terms of the simplicity and the accuracy in the evaluation of the off-diagonal elements.

One of the key concepts in developing the BOP algorithm is a change of basis. A bond-order expansion represented by a new basis can provide a turning point in the progress to the BOP, though the BOP formalism has not been derived from another basis instead of the atomic basis. Among various bases, I focus particularly on the Lanczos basis derived from the Lanczos algorithm for the following reasons: (1) the off-diagonal elements of the Green's function matrix are easily evaluated by the recurrence relation which follows from the tri-diagonalized Hamiltonian and the identity $(ZI-H)G(Z)=I$, and (2) the energy and forces on the central atom are calculated simultaneously at one Lanczos transformation.

In this paper I derive a BOP method based on the Lanczos basis within the two-center orthogonal TB model,[9] and show that the representation simplifies the evaluation of the forces. The method, which is an attempt to adopt another representation for the bond-order expansion, is a general recursion method to obtain the band energy and the forces. Moreover, as a test of the quality of this method, constant energy molecular dynamics (CEMD) simulations are carried out for carbon materials. The results of CEMD show that the BOP forces have sufficient accuracy in the molecular dynamics simulations.

Within the TB model,[9] the total potential energy $E_{\rm pot}$ of a system is a sum of band energy $E_{\rm band}$ and repulsive energy $E_{\rm rep}$, where $E_{\rm rep}$ is the sum of suitable repulsive pair potentials. The band energy in the BOP is described in two different terms of the LDOS and the bond order; these two quantities are not only formally identical, but also exactly identical in this presented BOP, as explained later. We can evaluate the force as the Hellmann-Feynman force from the latter rather than trying to differentiate the former.[3,4]

Assuming that the electrons are at a finite temperature $T$, the band energy $E_{\rm band}$ derived from the LDOS is expressed as follows:

$$E_{\rm band}=2\sum_{i\alpha}\int E n_{i\alpha}(E) f\left(\frac{E-\mu}{k_B T}\right) dE, \tag{1}$$

where $n_{i\alpha}(E)$ is the density of states projected onto $\alpha$ atomic orbital $|i\alpha\rangle$ in atom $i$, and the function $f(x)=1/[1+\exp(x)]$ is the Fermi function.

On the other hand, the band energy and the Hellmann-Feynman forces are written in terms of the bond orders $\theta$:

$$E_{\rm band}=\sum_{i\alpha,j\beta}\theta^{a}_{i\alpha,j\beta}H^{a}_{j\beta,i\alpha}, \tag{2a}$$

$$\mathbf{F}_k=-\sum_{i\alpha,j\beta}\theta^a_{i\alpha,j\beta}\frac{\partial H^a_{j\beta,i\alpha}}{\partial \mathbf{r}_k}, \tag{2b}$$

where $H^a_{j\beta,i\alpha}=\langle j\beta|\hat{H}|i\alpha\rangle$ is an element of the TB Hamiltonian matrix $H$, and the index $a$ indicates the representation based on the atomic basis.

Consider a change of representation, from the atomic basis into the Lanczos basis, by the block Lanczos algorithm.[10] The algorithm which is an efficient way for block tridiagonalizing a matrix is also the first step in calculating the diagonal elements of the Green's-function matrix. The central equation is

$$\hat{H}|U_n)=|U_n)\underline{A}_n+|U_{n-1})\underline{B}^\dagger_n+|U_{n+1})\underline{B}_{n+1}. \tag{3}$$

The $\underline{A}_n$ and $\underline{B}_n$ are recursion block coefficients with $p\times p$ in size, where $p$ is the number of atomic orbitals on the starting atom $i$. The states $|U_n)$ are constructed by the Lanczos bases, which are orthonormal and block tridiagonalize the Hamiltonian, as $|U_n)=(|L_{n1}\rangle,|L_{n2}\rangle,\ldots,|L_{np}\rangle)$. The representation based on atomic basis $T^a$ is transformed into that of the Lanczos basis $T^L$ by a matrix $U$:

$$T^L=U^\dagger T^a U, \tag{4}$$

where $U$ is defined by $U_{i\alpha,n\mu}=\langle i\alpha|L_{n\mu}\rangle$, and $T$ is the Hamiltonian $H$, the derivative of the Hamiltonian with respect to atomic coordinate $\partial H/\partial \mathbf{r}_i$, the bond order $\theta$, or the Green's-function $G(Z)$ matrix.

It is essential to choose a starting state $|U_0)$ in the block Lanczos algorithm as follows:

$$|U_0)=(|i1\rangle,|i2\rangle,\ldots,|ip\rangle). \tag{5}$$

Then considering Eq. (4) and the orthonormality of the Lanczos basis we have

$$\theta^a_{i\alpha,j\beta}=\sum_{n\mu}\theta^L_{0\alpha,n\mu}U^\dagger_{n\mu,j\beta}, \tag{6}$$

where $U^\dagger_{n\mu,j\beta}$ is a $(n\mu,j\beta)$ element in the $U^\dagger$ matrix. From Eq. (6) the bond orders based on the Lanczos basis are related to that by the atomic basis, which allows us to evaluate the bond order in the Lanczos representation. Let us introduce the block element $\underline{T}^L_{mn}=(U_m|\hat{T}|U_n)$, where $\hat{T}$ is an arbitrary operator, the size of the block element is $p\times p$, and the $(\alpha,\beta)$ element of $\underline{T}^L_{mn}$ is described as $T^L_{m\alpha,n\beta}$. Then the block elements which are needed for evaluating the bond orders based on the atomic basis in Eq. (6) are written as $\underline{\theta}^L_{0n}$. Therefore, starting the recursion with Eq. (5), we have only to evaluate the zeroth block line of the bond-order matrix. If the conventional scalar Lanczos algorithm [11–13] is applied for a change of representation, Eq. (6) is not simplified for other nonzero derivatives of the Hamiltonian with respect to atomic coordinate. Though it is also possible to start the recursion with a cluster containing a neighbor shell of atoms, the choice is not suitable for computer time.

The bond-order matrix can be related to the Green's function through the following equation:

$$\underline{\theta}^L_{0n}=-\frac{2}{\pi}\mathrm{Im}\left\{\int \underline{G}^L_{0n}(E+\mathrm{i}0^+)f\left(\frac{E-\mu}{k_BT}\right)dE\right\}, \tag{7}$$

where $0^+$ shows a positive infinitesimal. The integration of the Green's function with the Fermi function can be carried out in the complex plane by summing up an infinite series over the modified Matsubara poles.[3,4,14] This modified Matsubara summation converges rapidly with about 40 complex poles under a high electron temperature ($\gtrsim 1000$ K) though it often diverges under a lower electron temperature. The modified Matsubara summation is also employed for integrations of other Green's functions with the Fermi function.

The block element of the Green's function $\underline{G}^L_{00}(Z)$ is calculated by the recursion block coefficients $\underline{A}_n$ and $\underline{B}_n$ as a multiple inverse:

$$\underline{G}^L_{00}(Z)=[Z\underline{I}-\underline{A}_0-\underline{B}^\dagger_1[Z\underline{I}-\underline{A}_1-\underline{B}^\dagger_2[\cdots]^{-1}\underline{B}_2]^{-1}\underline{B}_1]^{-1}. \tag{8}$$

The $\underline{G}^L_{00}(Z)$ is equal to a block element represented by the atomic basis $\underline{G}^a_{kk}(Z)$, since we have started the block Lanczos algorithm with Eq. (5). Thus imaginary parts of Eq. (8) give the LDOS on atom $i$.

Moreover, by taking account of the block tridiagonalized Hamiltonian and the identity $(ZI-H)G(Z)=I$ in the Lanczos basis representation, the off-diagonal block elements of Green's function matrix $\underline{G}^L_{0n}(Z)$ are obtained from a recurrence relation:

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

where $\delta$ is the Kronecker's delta, and $\underline{G}^L_{0-1}$ and $\underline{B}^\dagger_0$ are $\underline{0}$, respectively. All the off-diagonal block elements $\underline{G}^L_{0n}(Z)$ are related to the diagonal block element $\underline{G}^L_{00}(Z)$. Once $\underline{G}^L_{00}(Z)$ has been obtained, the off-diagonal block elements are easily evaluated from the recursive calculation. The facility for the evaluation of the off-diagonal block elements is an important benefit produced by the Lanczos basis representation. Therefore the computational effort for the evaluation of the force is the same as that for the band energy. The diagonal element and the off-diagonal elements obtained from Eq. (8) and Eq. (9) are exact Green's functions in the recursion chain cluster block tridiagonalized with the Lanczos transformation, and the bond orders for the atom $i$ are evaluated through these Green's functions and Eq. (6). This means that Eq. (1) and Eq. (2a) are manifestly identical, since Eq. (2a) can be divided into contributions for each atom. On the other hand, Eq. (2b) is an approximate force for Eq. (1) or Eq. (2a). As the number of recursion levels increases, the Hellman-Feynman force Eq. (2b) converges the exact force.

In a series of the block Lanczos transformation, the Lanczos vectors hop from the central atom to the outside atoms, reflecting the local atomic environment. Since the Lanczos bases play the role of a perturbing medium for the central atom, for an infinite system, as the number of recursion levels in Eq. (8) increases, a good convergence is given for both band energy and the forces. However, we have to take only a finite number of levels for the recursion in Eq. (8) because of the computational effort, and then estimate the further levels. There are two simple ways of termination. [3] One is to take many levels on a cluster of small size without any terminator, and the other is to use a square root termi-

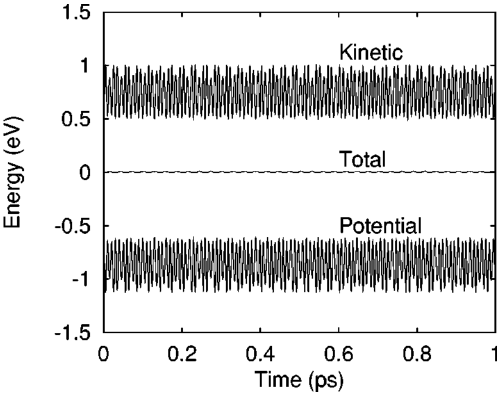

FIG. 1. The potential energy, kinetic energy, and total energy as a function of time for a constant energy molecular dynamics simulation of a carbon trimer.

nator which is derived assuming that the further block recursion coefficients are constant. I have adopted the former in this study.

In order to keep the number of electrons and to cancel the Madelung energy, though local charge neutrality (LCN) is often imposed as the simplest form of self-consistency in the usual BOP, the LCN is not an appropriate assumption with a physical meaning particularly for systems with the charge transfer. The LCN breaks even in a carbon trimer with the bond length of 1.3 Å , where the excess charge of a middle carbon is about 0.1. In this study, I have employed total charge neutrality (TCN) to perform molecular dynamics simulations while keeping the number of electrons. If the total excess charge on a system is $Q$, then a good estimate of the shift of the chemical potential is

$$\Delta=\lambda\frac{Q}{X}, \tag{10}$$

where $\lambda$ is a parameter to accelerate the convergence, and generally is 1.0. The total response function $X$ is given as follows:

$$X=-\frac{2}{\pi}\mathrm{Im}\left\{\sum_{i\alpha}\int[G^{a}_{i\alpha,i\alpha}(E+\mathrm{i}0^{+})]^{2}f\left(\frac{E-\mu}{k_BT}\right)dE\right\}. \tag{11}$$

Usually no more than three or four iterations are required to achieve the convergence that the absolute value of $Q/$atom is below $10^{-5}$. The TCN reduces the separability of individual atoms in the calculation of the band energy and the forces, and complicates slightly the parallelizability of the program code as compared with the LCN. However, the evaluation and the integration of the Green's function, which are time-consuming steps, are separately performed. Therefore the TCN retains the advantage of the BOP.

As a test of the consistency between the band energy and the forces, constant energy molecular dynamics simulations have been performed for a carbon trimer and diamond.[15] First, as an example of the finite system, a carbon trimer has been examined in CEMD using a time step of 0.5 fs. Here both Eqs. (8) and (9) are calculated with three recursion levels, the electron temperature $k_BT$ is 0.2 eV, 40 poles are used on the integration of the Green's function, and each atom is given an initial velocity corresponding to 1300 K. Figure 1 shows the energy of the carbon trimer as a function of time. The conserved total energy within $10^{-5}$ eV/atom indicates that the forces are of good quality.

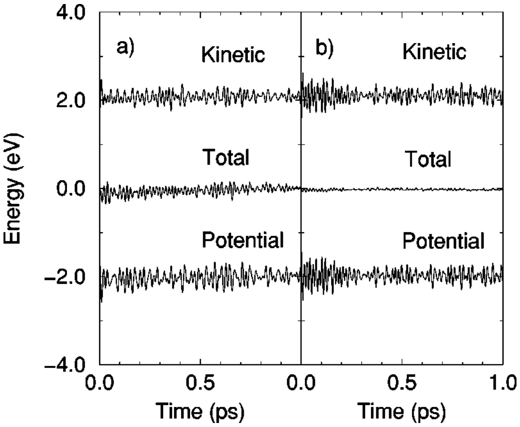

FIG. 2. The potential energy, kinetic energy, and total energy of a unit cell containing 64 atoms for constant energy molecular dynamics simulations of diamond at 300 K as a function of time. In (a) the results are for five recursion levels, and in (b) for ten recursion levels.

In a similar way, CEMD simulations have been carried out for diamond at 300 K using five and ten recursion levels on a small cluster without any terminator. The cutoff radius of the small cluster is 4.2 Å , which is sufficient for including the fifth neighbor atoms in a perfect lattice. In Fig. 2 the energy is shown for diamond in CEMD as a function of time. The fluctuation of the total energy decreases with increasing the number of recursion levels. For five recursion levels, the total energy shows an oscillation with amplitude of $10^{-2}$ eV/atom, while for ten recursion levels, the total energy is conserved within $10^{-3}$ eV/atom. The BOP with appropriate recursion levels provides an accurate band energy and forces in molecular dynamics simulations of a finite and infinite systems.

However, as the number of recursion levels increases, the computational effort becomes larger. Computer time and accuracy as a function of the number of recursion levels have been examined for the above CEMD of diamond on an IBM RS6000/SP2, where a cutoff radius was fixed at 4.2 Å , in all the number of recursion levels. Figure 3 shows the timing results for one time step and the amplitude for the total energy. The computer time is approximately proportional to the number of recursion levels, while the error for the total energy decreases rapidly.

Moreover, the method has been compared with direct diagonalization of the Hamiltonian with respect to computer time. Figure 4 shows the time for one step of the CEMD of diamond as a function of the number of atoms in the cell, for five levels, ten levels, and $k$ space ($\Gamma$). The computer time of the presented BOP method scales linearly with the number of atoms, and results in a smaller computer time than that of the diagonalization when the number of atoms exceeds about 100. Finally, I have performed parallel computation on a Sun Star Fire which is a parallel machine with 32 processors. It is observed that the scalability of the algorithm is almost ideal.

In the moment-based methods, it is well known that the vacancy formation energies of covalent materials such as carbon and silicon are not reproduced.[16,17] However, the pre-

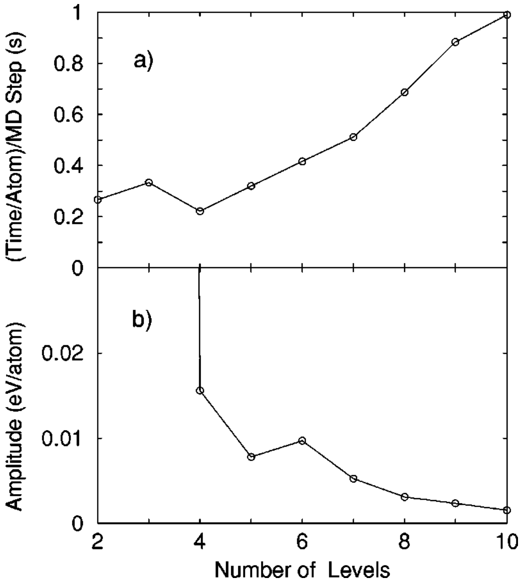

FIG. 3. (a) One atom energy and force calculation time. (b) The magnitude of total energy oscillation in diamond for constant energy molecular dynamics calculations as a function of the number of recursion levels (using a unit cell of 64 atoms at 300 K). The calculations were performed on an IBM RS6000/SP2 using one CPU.

sented BOP has reproduced the vacancy formation energy in preliminary calculations, which suggests that the method has an accuracy compared with the former moment-based methods. I plan to discuss details in the future.

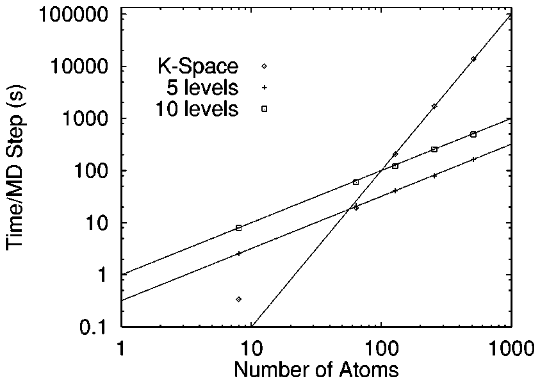

FIG. 4. Comparison between k space ($\Gamma$ point) and the BOP in terms of calculation time for the energy and force evaluation as a function of number of atoms in a unit cell for diamond. The calculations were carried out on an IBM RS6000/SP2 using one CPU.

In conclusion, I have presented a general recursion method, which is intrinsically linear in the scaling for large systems and very suitable for parallel computation, for calculating the band energy and the forces in tight-binding molecular dynamics simulations in terms of the bond orders based on the Lanczos basis. From practical applications with carbon materials, it has been demonstrated that the forces are of good quality.

I thank Y. Iwasa, T. Mitani, and N. Suzuki for encouragement and many enlightening discussions. I am grateful to M. Aoki for useful discussions and many comments.

---

[1] D. G. Pettifor, Phys. Rev. Lett. **63**, 2480 (1989).

[2] M. Aoki, Phys. Rev. Lett. **71**, 3842 (1993).

[3] A. P. Horsfield, A. M. Bratkovsky, D. G. Pettifor, and M. Aoki, Phys. Rev. B **53**, 1656 (1996).

[4] A. P. Horsfield, A. M. Bratkovsky, M. Fearn, D. G. Pettifor, and M. Aoki, Phys. Rev. B **53**, 12 694 (1996).

[5] G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992).

[6] X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10 891 (1993).

[7] M. S. Daw, Phys. Rev. B **47**, 10 895 (1993).

[8] S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994).

[9] V. Heine, Solid State Phys. **35**, 1 (1980).

[10] J. Inoue and Y. Ohta, J. Phys. C **20**, 1947 (1987).

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

[12] R. Haydock, V. Heine, and M. J. Kelly, J. Phys. C **5**, 2845 (1972); **8**, 2591 (1975).

[13] R. Haydock, Solid State Phys. **35**, 216 (1980).

[14] D. M. C. Nicholson, G. M. Stocks, Y. Wang, W. A. Shelton, Z. Szotek, and W. M. Temmerman, Phys. Rev. B **50**, 14 686 (1994).

[15] C. H. Xu, C. Z. Wang, C. T. Chan, and K. M. Ho, J. Phys.: Condens. Matter **4**, 6047 (1992).

[16] J. D. Kress and A. F. Voter, Phys. Rev. B **52**, 8766 (1995).

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