Tight-binding molecular dynamics with linear system-size scaling

1994 J. Phys.: Condens. Matter 6 9153

(http://iopscience.iop.org/0953-8984/6/43/015)

J. Phys.: Condens. Matter 6 (1994) 9153–9172. Printed in the UK

Tight-binding molecular dynamics with linear system-size scaling

S-Y Qiu, C Z Wang, K M Ho and C T Chan
Ames Laboratory and Department of Physics, Iowa State University, Ames, Iowa 50011, USA

Received 26 May 1994, in final form 22 August 1994

Abstract. A tight-binding molecular dynamics (TBMD) scheme with linear system-size scaling is implemented by incorporating the density-matrix electronic-structure method into tight-binding molecular dynamics. We demonstrate that this scheme, compared to the TBMD with standard diagonalization methods, can work more efficiently for systems larger than a few hundred atoms. We present our testing results on crystalline, amorphous and liquid carbon systems in order to establish the general applicability of the scheme to systems under various physical conditions.

1. Introduction

Recently, many research efforts on electronic structure calculations have been concentrated on the development of algorithms with linear system-size scaling (also called order-N algorithms) [1–9]. This is motivated by the fact that standard algorithms have a complexity that grows as the cube of the system size, and thus the use of first-principles and tight-binding calculations are limited to small systems. Therefore, an order-N algorithm, if it is efficient and accurate, will evidently broaden the application of first-principles and tight-binding calculations to a wide variety of new problems involving larger and more complicated systems. An additional advantage is that order-N algorithms tend to be more naturally adaptable to parallel computers. With an order-N scheme implemented on parallel machines it is expected that many simulations which were too expensive to be carried out previously can now be performed fairly quickly.

The density-matrix (DM) algorithm developed by Li and co-workers [1] and by Daw [2] is one of the most promising order-N algorithms for electronic structure calculations. In their approach, they introduce a variational method for solving the electron density matrix. The method takes advantage of the locality of the density matrix in real space to achieve linear scaling. The approximation is made by truncating the off-diagonal elements of the density matrix that correspond to atomic pairs whose distances are beyond a cutoff radius R_c. The method becomes exact as $R_c \to \infty$. The solution of the variational problem involves only an unconstrained minimization, which may be performed by conjugate-gradient or other standard techniques. This is very well suited for molecular dynamics (MD) simulations.

In the past several years, we have developed a molecular-dynamics scheme [10] that incorporates electronic structure effects into MD simulations through a tight-binding parametrization. In this TBMD approach, the covalent bonding of the material enters the calculations in a natural way from the underlying electronic structure, rather than through an *ad hoc* N-body potential. Unlike the Car–Parrinello approach, which relies on the expansion of the electronic wavefunctions by plane waves [11], the tight-binding electronic calculations require only a few atomic orbitals for each atom, allowing a larger number

0953-8984/94/439153+20$19.50

of atoms and longer simulation periods to be tackled within a given computer capability. However, the bottleneck of the TBMD calculation is the direct-diagonalization (DD) of the tight-binding Hamiltonian, which scales as the cube of the system size, and thus limits the simulation to small systems.

In this paper, we report results of incorporating the DM method into the TBMD to develop an order-N TBMD scheme (DM-TBMD). We test the scheme on carbon systems, where a well tested tight-binding Hamiltonian has already been developed [12]. The results obtained from the DM-TBMD scheme are compared with those from the direct-diagonalization method (DD-TBMD). We first study the energy against volume curves for various coordinated crystalline structures of carbon. The purpose is to find an acceptable minimum cutoff radius R_c for the electron density matrix, such that the approximated results are accurate enough to compare with the exact eigensolution. Then, we test the DM-TBMD scheme on various carbon systems, crystalline, amorphous and liquid carbon, in order to establish the general applicability of the scheme to systems under various physical conditions. Finally, we discuss the efficiency of the scheme and measure the crossover point at which DM-TBMD becomes more efficient than DD-TBMD.

2. DM-TBMD

In the TBMD scheme [10, 13–15], the system is described by a Hamiltonian of the form

$$H(\{r_i\}) = \sum_i \frac{P_i^2}{2m} + \sum_n^{\text{occupied}} \langle \psi_n | H_{\text{TB}}(\{r_i\}) | \psi_n \rangle + E_{\text{rep}}(\{r_i\}) \tag{2.1}$$

where $\{r_i\}$ denotes the positions of the atoms ($i = 1, 2, \ldots, N$) and P_i denotes the momentum of the ith atom. The first term in (2.1) is the kinetic energy of the ions, the second term is the electronic band-structure energy, $E_{\text{TB}}(\{r_i\})$, calculated from a parametrized tight-binding Hamiltonian $H_{\text{TB}}(\{r_i\})$, and the third term is a short-ranged repulsive energy.

The tight-binding model used in the present simulation consists of an orthogonal sp^3 basis with on-site atomic energies $\epsilon_s = -2.99\,\text{eV}$ and $\epsilon_p = 3.71\,\text{eV}$ and two-centre integrals

$$V_{ss\sigma}(r) = -5.00h(r)\,\text{eV} \qquad V_{sp\sigma}(r) = 4.70h(r)\,\text{eV}$$
$$V_{pp\sigma}(r) = 5.50h(r)\,\text{eV} \qquad V_{pp\pi}(r) = -1.55h(r)\,\text{eV}$$

where $h(r)$ is a smooth function of interatomic distance. The repulsive energy is in the form of $E_{\text{rep}} = \sum_i f[\sum_j \phi(r_{ij})]$, where $\phi(r_{ij})$ is a pairwise repulsive interaction and f is a functional with argument $x = \sum_j \phi(r_{ij})$. More details about the tight-binding model can be found in [12]. The accuracy and transferability of this model have been well tested. It reproduces well the first-principles energy against volume curves for various coordinated crystalline structures of carbon. In particular, the energy curves of the linear chain (twofold), graphite (threefold) and diamond (fourfold) structures are excellently described [12, 16]. The model also describes well the elastic, vibrational and anharmonic properties of diamond and graphite [12], as well as the properties of more complex systems such as carbon clusters with sizes ranging from 5–100 atoms [17].

Instead of calculating E_{TB} through the direct-diagonalization of the H_{TB}, which scales as the cube of the system size, the DM method developed by Li and co-workers [1] performs the energy calculation through the following variational approach. The grand potential defined as

$$\Omega = E_{\text{TB}} - \mu N_e = \text{tr}[\tilde{\rho}(H_{\text{TB}} - \mu)] \tag{2.2}$$

is minimized with respect to the density matrix $\tilde{\rho}$, where N_e is the total number of valence electrons in the system and μ is the chemical potential. The density matrix $\tilde{\rho}$ is related to a variational matrix ρ through

$$\tilde{\rho} = 3\rho^2 - 2\rho^3 \tag{2.3}$$

where off-diagonal elements of ρ beyond a cutoff range are set to zero by assuming that ρ is well localized. In the appendix we give a general discussion of the density matrix scheme in comparison with the other order-N schemes proposed in [3–5].

The minimization of Ω can be carried out either with conjugate-gradient or steepest-descent algorithms. Because of the truncation of ρ, the calculated total number of valence electrons, $N_e = \mathrm{tr}[\tilde{\rho}]$, will usually not be exact, even though the chemical potential μ is fixed at the gap between the valence and conduction bands. If a system has a relatively large gap, such as for a system of insulators or semiconductors, the error in the calculated N_e can be kept small with a constant input μ fixed at the gap. However, for metallic systems, it is difficult to have a convergence of N_e while fixing μ in the simulation. In order to have a converged N_e, μ also needs to be calculated iteratively at each MD step. Here we propose a two-stage steepest-descent minimization algorithm, where the chemical potential μ is automatically adjusted to the correct value by imposing the condition that N_e is equal to the total number of valence electrons.

At each iteration of steepest-descent minimization, our algorithm for the line minimization can be divided into two stages. In the first stage, the line minimization is carried out along the direction of $-\nabla_\rho \Omega|_{\rho=\rho_n} \equiv A_n$, where explicitly

$$A_n = \left[-\nabla_\rho E_{\mathrm{TB}}|_n\right] - \mu_n\left[-\nabla_\rho N_e|_n\right] \tag{2.4}$$

$$= \left[-\nabla_\rho \mathrm{tr}[\tilde{\rho} H_{\mathrm{TB}}]|_{\rho=\rho_n}\right] - \mu_n\left[-\nabla_\rho \mathrm{tr}[\tilde{\rho}]|_{\rho=\rho_n}\right] \tag{2.5}$$

where $\tilde{\rho}_n$ and μ_n are the density matrix and the chemical potential at the nth iteration, respectively. Along that search direction with step-size λ, we have the variational matrix at the $(n+1)$th iteration after the first-stage line minimization

$$(\rho_1)_{n+1} = \rho_n + \lambda A_n \tag{2.6}$$

where the value of λ can be found by putting (2.6) into (2.2), and thus obtaining the grand potential as a third-order polynomial of λ

$$(\Omega_1)_{n+1} = c_0 + c_1\lambda + c_2\lambda^2 + c_3\lambda^3 \tag{2.7}$$

where the four coefficients are related to the parameters at the nth iteration through the following equations:

$$c_0 = \Omega_n \tag{2.8}$$

$$c_1 = -\mathrm{tr}[A_n^2] \tag{2.9}$$

$$c_2 = \mathrm{tr}[3A_n^2 H - 2A_n^2(H\rho_n + \rho_n H) - 2A_n\rho_n A_n H] \tag{2.10}$$

$$c_3 = -2\mathrm{tr}[A_n^3 H] \tag{2.11}$$

where $H \equiv H_{\mathrm{TB}}$. Now the value of λ (denoted as λ_{min}) at which $(\Omega_1)_{n+1}$ is at the minimum can be easily calculated.

Since after the first-stage line minimization the number of electrons calculated by $N_e = \mathrm{tr}[(\tilde{\rho}_1)_{n+1}]$ will generally not be equal to the actual number of valence electrons in the system, we now adjust the chemical potential μ at the second stage of the line minimization to eliminate the electron number discrepancy. We adjust μ along the direction of $-\nabla_\rho N_e|_{\rho=\rho_n} \equiv B_n$ with step-size $\delta\mu$; the density matrix at the $(n+1)$th iteration after the second-stage line minimization is then

$$(\rho_2)_{n+1} = \rho_n + \lambda_{\min} A_n + \delta\mu B_n \tag{2.12}$$

where the value of $\delta\mu$ can be found by putting $(\rho_2)_{n+1}$ into the equation $\mathrm{tr}[(\tilde{\rho}_2)_{n+1}] = N_e^{\mathrm{exact}}$ to form a cubic equation of $\delta\mu$:

$$d_0 + d_1\delta\mu + d_2\delta\mu^2 + d_3\delta\mu^3 = N_e^{\mathrm{exact}} \tag{2.13}$$

where N_e^{exact} is the exact value of the total number of valence electrons, and the four coefficients are related to the parameters at the nth iterations through the following equations:

$$d_0 = \mathrm{tr}[(\tilde{\rho}_1)_{n+1}] \tag{2.14}$$

$$d_1 = -\mathrm{tr}[B_n(B_1)_{n+1}] \tag{2.15}$$

$$d_2 = 3\mathrm{tr}[B_n^2 - 2B_n^2(\rho_1)_{n+1}] \tag{2.16}$$

$$d_3 = -2\mathrm{tr}[B_n^3] \tag{2.17}$$

where $(B_1)_{n+1} = -\nabla_\rho N_e|_{\rho=(\rho_1)_{n+1}}$. Among the three roots of the cubic equation, we find it convenient to use the physical root with the smallest absolute value as the solution (denoted as $\delta\mu_{\min}$). Hence, by combining the above two stages, we obtain the variational matrix ρ_{n+1} at the $(n+1)$th iteration to be

$$\rho_{n+1} = \rho_n + \lambda_{\min} A_n + \delta\mu_{\min} B_n \tag{2.18}$$

and the new chemical potential at the $(n+1)$th iteration is

$$\mu_{n+1} = \mu_n - \frac{\delta\mu_{\min}}{\lambda_{\min}}. \tag{2.19}$$

In summary, for the above two-stage steepest-descent minimization algorithm, we first minimize Ω along the direction of $-\nabla_\rho\Omega$, and then adjust μ along the direction of $-\nabla_\rho N_e$, so that the density matrix at each iteration is always on the surface of $\mathrm{tr}[\tilde{\rho}] = N_e^{\mathrm{exact}}$.

The Hellmann–Feynman force can be calculated through a derivative of the grand potential Ω with respect to a parameter ξ (an atomic coordinate, for example) at fixed μ:

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\xi} = \frac{\partial\Omega}{\partial\rho}\frac{\mathrm{d}\rho}{\mathrm{d}\xi} + \frac{\partial\Omega}{\partial H_{\mathrm{TB}}}\frac{\mathrm{d}H_{\mathrm{TB}}}{\mathrm{d}\xi} \tag{2.20}$$

but the first term vanishes at the variational solution, so that the force is given by

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\xi} = \mathrm{tr}\left[\tilde{\rho}\frac{\mathrm{d}H_{\mathrm{TB}}}{\mathrm{d}\xi}\right]. \tag{2.21}$$

The DM method calculates *approximately* the electron band-structure energy E_{TB} and the corresponding force in an approach that scales linearly with the system size. The approximation becomes exact when no truncation is performed on the variational matrix ρ. Further details about the DM method can be found in [1, 2].

3. Simulation and results

3.1. Energy against volume curves

Using the DM method, we first need to reproduce the energy against volume curves for various crystalline structures of carbon calculated by direct diagonalization of the tight-binding Hamiltonian. In order to obtain an order-N scheme, an appropriate cutoff radius R_c has to be chosen for the variational density matrix ρ. However, unless a large enough R_c is chosen, or R_c is scaled with the volume, a constant cutoff radius will give rise to discontinuities in the energy against volume curves. This is because, for a constant R_c, the number of atoms that 'see' a particular atomic site is larger for a smaller volume than for a larger volume, and thus energies are calculated with more accuracy in the former case. In MD simulations at high temperature, such energy discontinuities can also occur because diffusion of the atoms can bring them into or out of R_c. Since choosing a large R_c is inefficient, and varying R_c is impractical in MD simulations, we shall use a constant number N_c, the number of closest neighbours for each atom, whose corresponding off-diagonal elements of the variational density matrix are non-zero. That is, for each atom, we index all other atoms in incremental order of their distance to that central atomic site. We then choose a cutoff number N_c, and for the first N_c neighbours in the list we choose their corresponding off-diagonal elements of the matrix ρ to be variational parameters, and set all the other off-diagonal elements to zero. In this way, the same accuracy can be achieved for both small and large volumes, and a smooth energy against volume curve can be obtained.

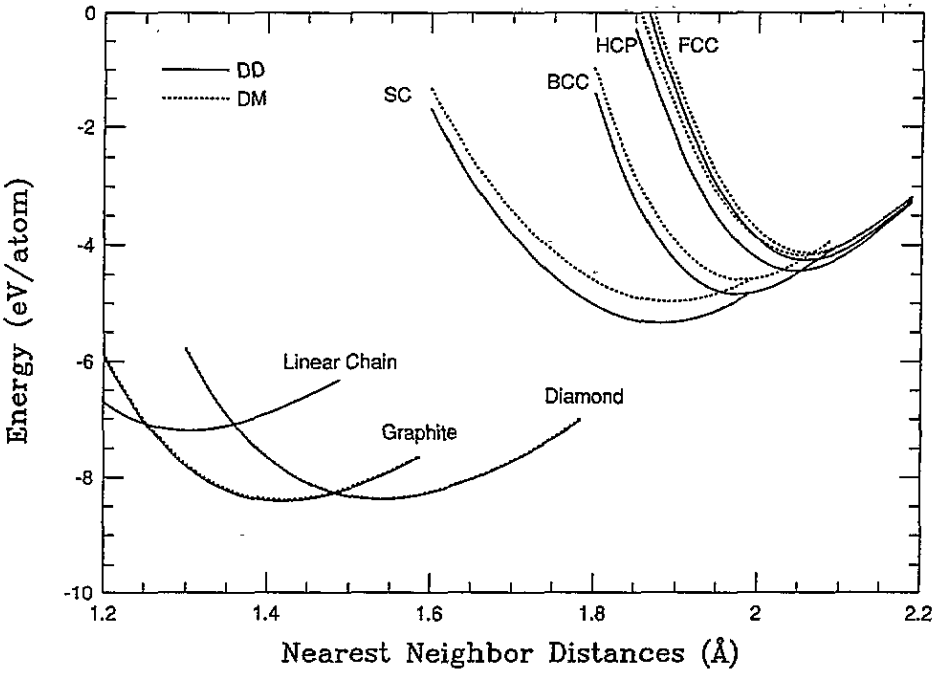

Figure 1. Cohesive energy against nearest-neighbour atomic distance curves of the diamond, graphite, linear chain, SC, BCC, HCP and FCC structures. Full and broken curves show the results from DD and DM calculations, respectively.

In figure 1 we show energy against volume curves for various crystalline structures of carbon: diamond (64, 43), graphite (80, 44), linear chain (80, 46), simple cubic (SC) (64, 41), face-centre cubic (FCC) (64, 45) body-centre cubic (BCC) (54, 46) and HCP (72, 43) where, in the parentheses, the first number is the number of atoms per unit cell, and the second number is the cutoff neighbours N_c, which include complete atomic shells of crystalline structures. Only the Γ point is used for the electronic structure calculation, and a cubic periodic boundary condition is imposed for each system. As can be seen from figure 1, the DM method at those cutoff neighbours N_c accurately reproduces the energy against volume curves calculated by the DD method, especially for the diamond, graphite and linear chain cases. The relatively poor fit for the SC, FCC, BCC and HCP systems is due to the fact that these are metallic systems, and thus a larger cutoff N_c is required in order to achieve the same good results as for the diamond, graphite or linear chain systems.

In what follows, we perform a DM-TBMD simulation on carbon using various system sizes and densities. Based on the above calculation, we shall use a constant N_c (= 46) for the variational matrix ρ cutoff. In the MD simulation, the energy discontinuity will not be completely eliminated. This is because, during the simulation, not only the actual atoms included in the list of the first N_c neighbours will vary during the simulation, the actual number N_c for each atom cannot be kept the same. The first problem will appear when atoms diffuse, and the second problem is due to the symmetry requirement of the density matrix. For example, suppose atom 2 is in the list of the first N_c neighbours of atom 1, while atom 1 is not in the list of the first N_c neighbours of atom 2, which means that the 4×4 matrix block ρ_{12} is non-zero while ρ_{21} is zero. Since ρ is symmetric, atom 1 has to be added into the neighbour list of atom 2 to make ρ_{21} non-zero; thus atom 2 will have $N_c + 1$ neighbours in its list. This example tells us that, during the MD simulation, at best we can only use the same N_c for each atom not as the actual size but as the minimum size of its neighbour list. We observed in our calculation that this constant-minimum-N_c cutoff is still superior to the constant-R_c cutoff as far as energy conservation is concerned.

3.2. Crystalline carbon

Using the DM-TBMD method, as described in the previous section, we first performed a simulation on the ideal diamond structure using 64 and 216 atom unit cells of carbon with cubic periodic boundary conditions. Only the Γ point is used for the electronic structure calculation. Our simulation was initiated with all atoms arranged in the ideal diamond lattice and giving each atom a random distortion, which is equivalent to a temperature of about 400 K. The equations of atomic motions were solved by a fifth-order predictor–corrector algorithm with a timestep of 0.7×10^{-15} s. The minimum N_c for each atom is set to 46. We use the steepest-descent algorithm for the minimization of Ω in (2.2), with the chemical potential μ fixed at 3.0 eV in the middle of the gap between the valence and the conduction band. The minimization tolerance τ is set to 10^{-5}, where τ is defined as $\tau = |\Delta\Omega_n/\Omega_n|$, and $\Delta\Omega_n = \Omega_n - \Omega_{n-1}$, the difference of Ω between the nth and $(n-1)$th steepest-descent iterations. We found that the total number of valence electron $\mathrm{tr}[\tilde{\rho}]$ is quite close to the actual value at the present cutoff, and the error can be kept small in the entire simulation, as shown in figure 2(*d*).

The initial variational electron density matrix ρ at the first MD step can start from scratch, as suggested by Li and co-workers [1]: 0.5 for diagonal elements and zero for off-diagonal elements. In order to speed up the convergence of the steepest-descent minimization for the subsequent MD timesteps, the initial density matrix can be predicted by extrapolating forward from the electronic configurations of previous timesteps. A second-order extrapolation is

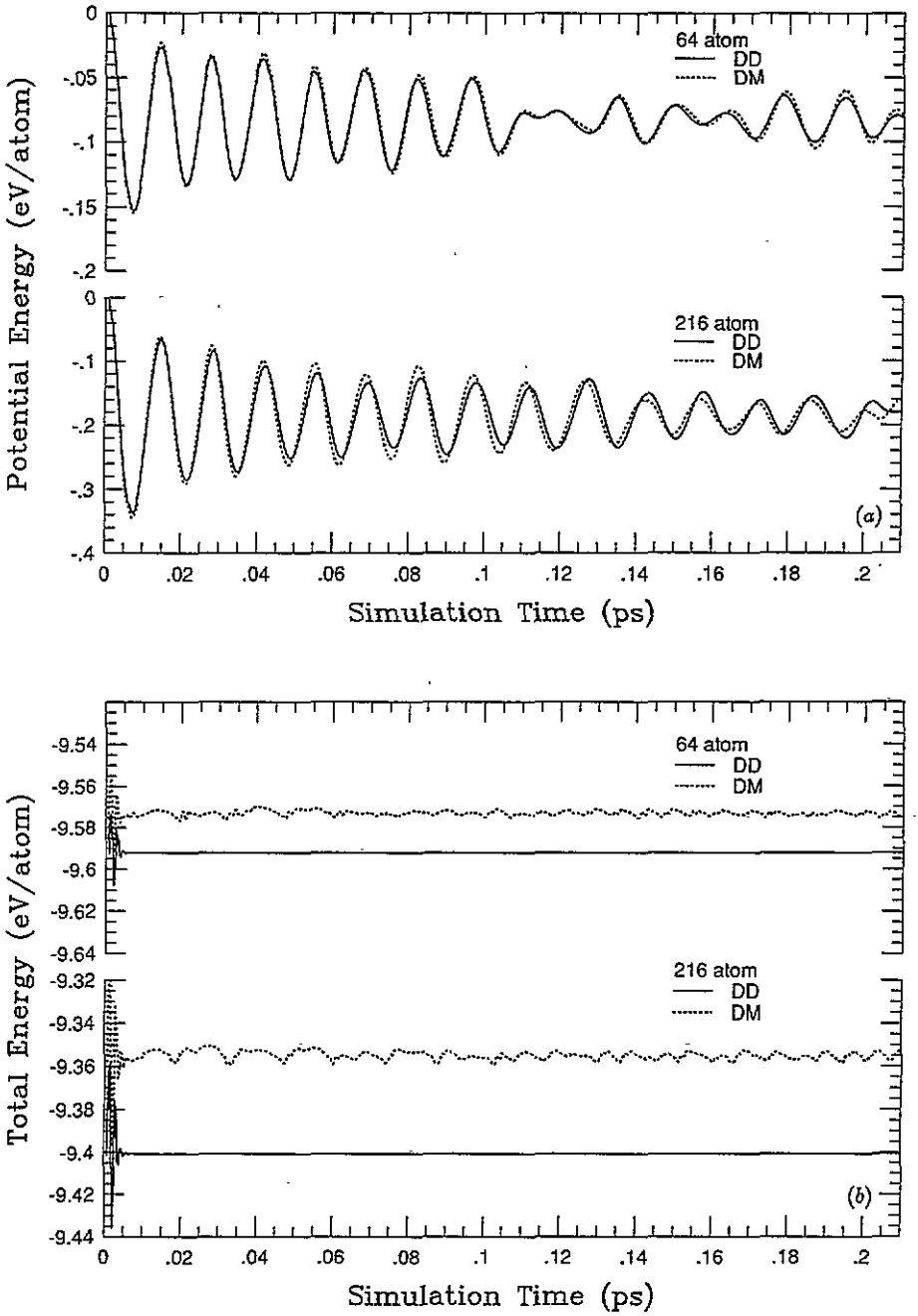

Figure 2. TBMD simulation for crystalline carbon for the period 0.21 ps (300 MD timesteps) with conjugate gradient tolerance $\tau = 10^{-5}$ for system sizes of 64 and 216 carbon atoms: (*a*) potential energy and (*b*) system total energy.

defined as

$$\rho'(\{r(t_{n+1})\}) = \rho(\{r(t_n)\}) + \alpha[\rho(\{r(t_n)\}) - \rho(\{r(t_{n-1})\})] + \beta[\rho(\{r(t_{n-1})\}) - \rho(\{r(t_{n-2})\})] \tag{3.1}$$

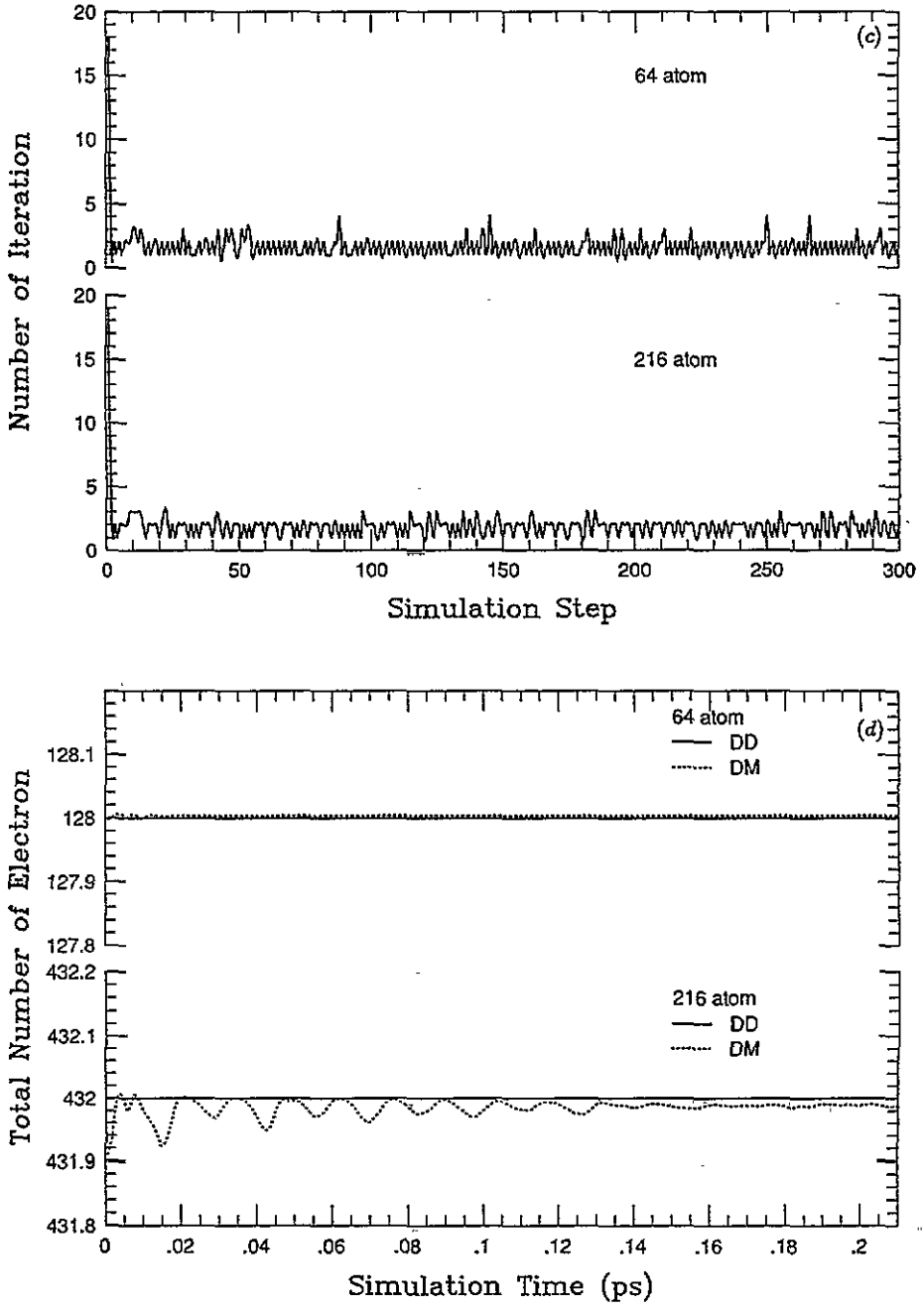

Figure 2. (Continued.) TBMD simulation for crystalline carbon for the period 0.21 ps (300 MD timesteps) with conjugate gradient tolerance $\tau = 10^{-5}$ for system sizes of 64 and 216 carbon atoms: (*c*) the number of conjugate gradient iterations and (*d*) total number of valence electrons.

where $\{r(t_n)\}$ are the ionic coordinates at time t_n, with n being the number of MD timesteps, α and β are fitted parameters, and the prime indicates an initial trial density matrix, as opposed to the fully converged density matrix $\rho(\{r(t_{n+1}\})$. This is inspired by the work of Payne and

co-workers [18] and Arias and co-workers [19]: if the initial electronic configuration can be moved closer to the correct instantaneous ground-state configuration, less computational effort is required to converge it to its exact ground state.

Although the parameters α and β can be found by minimizing the difference of the trial ionic coordinates

$$\{r'(t_{n+1})\} = \{r(t_n) + \alpha[r(t_n) - r(t_{n-1})] + \beta[r(t_{n-1}) - r(t_{n-2})]\} \tag{3.2}$$

with the actual ionic locations $\{r(t_{n+1})\}$ as suggested by Arias and co-workers [19], we have found that it is always more efficient to simply choose $\alpha = 2$ and $\beta = -1$. This is arrived at by considering the Taylor expansion of $\{r(t_{n-1})\}$ and $\{r(t_{n-2})\}$ around $\{r(t_n)\}$:

$$\{r(t_{n-1})\} = \{r(t_n)\} - \frac{\mathrm{d}\{r(t_n)\}}{\mathrm{d}t}\Delta t + \frac{1}{2!}\frac{\mathrm{d}^2\{r(t_n)\}}{\mathrm{d}t^2}\Delta t^2 - \cdots \tag{3.3}$$

$$\{r(t_{n-2})\} = \{r(t_n)\} - \frac{\mathrm{d}\{r(t_n)\}}{\mathrm{d}t}(2\Delta t) + \frac{1}{2!}\frac{\mathrm{d}^2\{r(t_n)\}}{\mathrm{d}t^2}(2\Delta t)^2 - \cdots. \tag{3.4}$$

By substituting (3.3) and (3.4) into (3.2) and including only up to second order, we obtained the above result. Similarly, for the first-order extrapolation, we can simply choose $\alpha = 1$ and $\beta = 0$, and for the zero-order $\alpha = \beta = 0$.

For the DM-TBMD simulation on crystalline carbon we have found that, if we use first-order or zero-order extrapolations, the total energy of the system constantly leaks. This is due to the accumulation of systematic errors in the steepest-descent minimization, since for any finite tolerance the first term of (2.10), $\partial\Omega/\partial\rho$, is not exactly zero. It is possible to conserve the total energy to a very high degree of precision either with a sufficiently small tolerance for the steepest-descent minimization ($\tau < 10^{-10}$), or by starting from scratch at each MD timestep in order to eliminate the error accumulation. Although the latter approach still gives an error in the force calculation depending on the choice of tolerance, the error tends to be random rather than systematic, so that the energy fluctuates rather than leaks constantly. However, we cannot reach maximum efficiency with either method. Nevertheless, for this particular system of crystalline carbon, we found that the system energy increases monotonically if we use the second-order extrapolation, while a first-order extrapolation causes the energy to decrease monotonically. Thus, by alternately using first-order and second-order extrapolations, we were able to achieve conservation of the total energy of the system using a relatively large tolerance for the minimization process ($\tau = 10^{-5}$).

To demonstrate the quality of the results we show, in figure 2(*a*), the evolution of potential energy ($E_{\mathrm{TB}} + E_{\mathrm{rep}}$), and in (*b*) the total energy of the system, against simulation time for a tolerance of 10^{-5}. As can be seen in figure 2(*a*), the evolution of potential energy calculated from the DM-TBMD method agrees remarkably well with that from the DD-TBMD. Although the system total energy calculated from DM-TBMD is not constant, it fluctuates around a constant value, which is acceptable for MD simulations. In figure 2(*c*), we show the number of steepest-descent iterations against simulation steps. As can be seen, except for the first step, which takes eighteen iterations to converge from scratch, the average iteration number for subsequent timesteps is around two. The error for the total number of valence electrons is small, less than 0.003% for the system of 64 atoms and less than 0.02% for 216 atoms, as shown in figure 2(*d*). The crossover point, at which the DM-TBMD scheme becomes more efficient than DD-TBMD, is around 60 atoms, as shown in figure 3.

Here, we report the timing of our program measured on an IBM 6000/560 workstation, for the purpose of performance comparison with the implementation of the density matrix

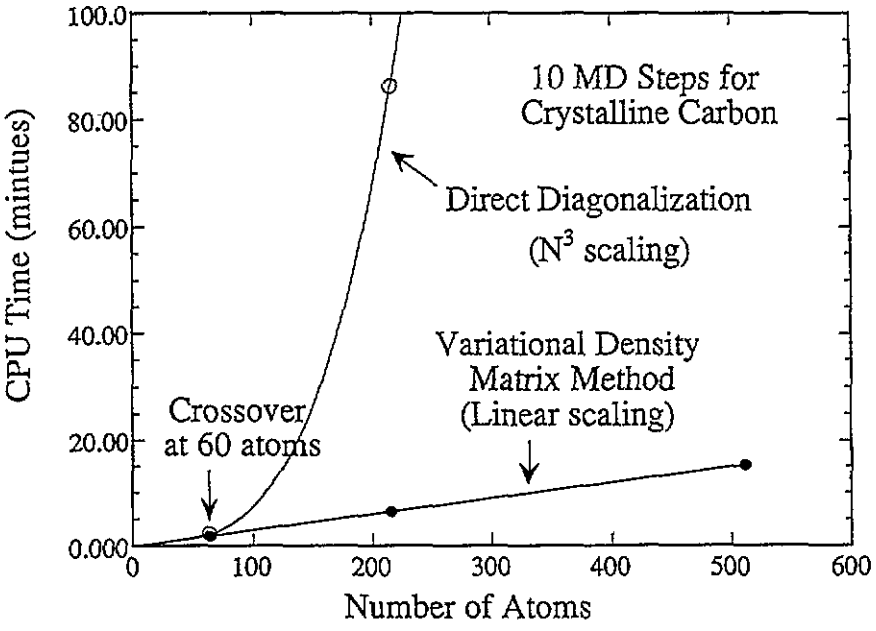

Figure 3. The measure of crossover point between the DM-TBMD and DD-TBMD method. The data are collected from the simulation of crystalline carbon over 10 MD timesteps for a system size of 64, 216 and 512 carbon atoms per unit cell.

method by other groups. If we use a constant λ in the steepest-descent minimization, we reach a speed of 0.056 (second per atom per iteration). If we calculate optimal λ_{min} at each steepest-descent minimization iteration, we reach a speed of 0.090 (second per atom per iteration), which is the calculation we performed for the crystalline carbon system. Finally, if we calculate both λ_{min} and $\delta\mu_{min}$ at each steepest-descent minimization iteration as described in (2.18), we obtain a speed of 0.110 (second per atom per iteration), which is the calculation we shall perform in the following for the amorphous and liquid carbon systems.

3.3. Amorphous carbon

After the success of the DM-TBMD technique on crystalline carbon, as shown in the previous section, we applied the scheme to a more disordered system, an amorphous carbon structure. The amorphous carbon is generated by quenching high-density high-temperature liquid carbon through TBMD simulations [20]. It has a density of $3.50\,g\,cm^{-3}$ with 216 carbon atoms per unit cell and cubic periodic boundary conditions imposed. The DM-TBMD simulation is performed in the same way as for the crystalline case, except that (i) the temperature is now 700 K, (ii) the timestep is 1.05×10^{-15}, (iii) we use the two-stage steepest-descent minimization to adjust the chemical potential μ at each step in order to obtain the actual number of valence electron, and (iv) the tolerance for the steepest-descent minimization is set to 10^{-6}. The parameters in (i) and (ii) are chosen to be the same as in [20]. The two-stage steepest-descent algorithm is used here because, if we use fixed μ in the simulation, we cannot obtain a well converged N_e as in the crystalline carbon case.

As can be seen in figure 4, the DM-TBMD scheme also works quite well for amorphous carbon. Though the potential energy against simulation time does not match the exact

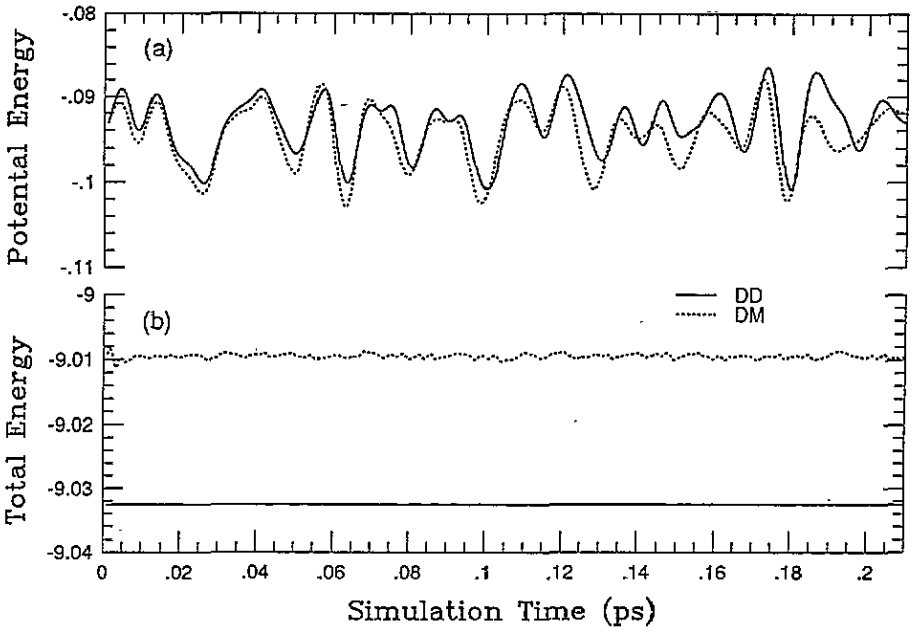

Figure 4. TBMD simulation for amorphous carbon for the period of 0.21 ps (200 MD timesteps) with conjugate gradient tolerance $\tau = 10^{-5}$ for system sizes of 216 carbon atoms: (*a*) potential energy (eV per atom) and (*b*) system total energy (eV per atom).

solution as well as the crystalline carbon case, it can be improved by increasing the cutoff range, improving the minimization tolerance and reducing the MD timestep. The total energy of the system fluctuates around a constant value, which is conserved quite well in the entire simulation. The average number of iterations for the steepest-descent minimization is around seven, and the crossover point is around 120 atoms.

3.4. Liquid carbon structure

In our previous work using DD-TBMD [16], we performed extensive simulations on liquid carbon and showed that at low density ($2.0\,g\,cm^{-3}$) the liquid carbon is metallic and dominated by twofold- and threefold-coordinated atoms. In this section, we want to demonstrate that the DM-TBMD scheme can closely reproduce the DD-TBMD results for liquid carbon, a more disordered system than both crystalline and amorphous carbon.

We performed the simulations with a density of $2.0\,g\,cm^{-3}$ using 64 atoms per unit cell and a cubic periodic boundary condition. Only the Γ point is used for the electronic structure calculation, and the cutoff neighbour number N_c for the variational density matrix is set to 46. The equations of atomic motions were solved by the fifth-order predictor–corrector algorithm with a timestep of 0.7×10^{-15}. After 3.5 ps of thermalization at 5000 K, we released the temperature control and ran another 1.4 ps using both DM-TBMD and DD-TBMD schemes. For metallic systems, it is difficult to guess μ correctly to give the correct total number of valence electrons; we use the two-stage steepest-descent minimization algorithm with μ adjusted at each iteration to obtain the actual number of valence electrons. The results of pair-correlation functions and atomic distributions obtained from the DM-TBMD and DD-TBMD schemes are compared in figure 5(*a*) and table 1. These results show that the DM-TBMD scheme reproduces quite well the results of DD-TBMD. To further demonstrate

the agreement of results from the two schemes, we have plotted in figures 5(*b*), 5(*c*) and 5(*d*) the partial radial distribution functions and bond-angle distribution functions of various coordinated atoms, respectively.

Since liquid carbon is a more disordered system than crystalline and amorphous carbon structures, the evolution of system potential energy calculated from DM-TBMD cannot match what is obtained from the DD-TBMD, as can be seen in figure 6. In fact, their trajectories in phase space begin to diverge at about 50 MD timesteps. This is not surprising, since for chaotic dynamics any two schemes with a slight difference will lead to a divergence over a long enough time period. This can be clearly seen in figure 7(*a*), where the liquid carbon system is slightly distorted (broken curve), and its trajectory using the same DD-TBMD scheme begins to diverge from the original one at around 60 MD steps. For comparison, we also show in figure 7(*b*) the corresponding crystal carbon case, where the same amount of distortion is made, and its trajectory can still follow quite well, as discussed previously. Moreover, as shown in figure 6 (thick dotted curve), if we interrupt the DM-TBMD at the 150th timestep and then continue it with the ionic configuration from DD-TBMD at that point, the results from the two schemes begin to diverge again after about 50 MD timesteps of quite good agreement. Nevertheless, we showed that this divergence does not affect the time-averaged physical properties of the system. In order to achieve conservation of the total energy of the system we found that, if we use the extrapolation technique described above for the prediction of the trial density matrix for the next MD timestep, we have to use a very small tolerance for the steepest-descent minimization ($\tau = 10^{-10}$). However we also found that, if we simply use the scratch density matrix for the starting point for each MD timestep, energy conservation can be achieved fairly well with a relatively large tolerance of $\tau = 10^{-4}$ (the average iteration number is around 25), which turns out to be more efficient. The above results are obtained using the latter approach. The crossover point is around 230 atoms.

Table 1. Ratios of various coordinated atoms of liquid carbon under density 2.0 g cm^{-3} calculated from DD-TBMD and DM-TBMD; n_c is the average coordination number.

Scheme	Onefold (%)	Twofold (%)	Threefold (%)	Fourfold (%)	n_c
DD-TBMD	3.85	40.85	50.75	4.54	2.56
DM-TBMD	3.62	42.63	49.59	4.15	2.54

4. Discussion and conclusion

In this paper we have demonstrated that the DM-TBMD scheme with linear system-size scaling works remarkably well for various carbon systems. For crystalline and amorphous carbon structures, the DM-TBMD can reproduce well the evolution of system potential energy calculated from the DD-TBMD. For liquid carbon, the DM-TBMD scheme can reproduce the atomic coordination number and pair-correlation functions from the DD-TBMD.

Efficiency and accuracy are always a trade-off that one has to balance. In the DM-TBMD scheme this trade-off is related to two parameters: the cutoff range for the variational density matrix and the tolerance for the steepest-descent minimization. A large cutoff range and small tolerance give high accuracy but low efficiency. Thus one should choose the smallest cutoff range and the largest tolerance to produce the most efficient MD simulation while still being able to obtain the same physics as the DD-TBMD scheme.

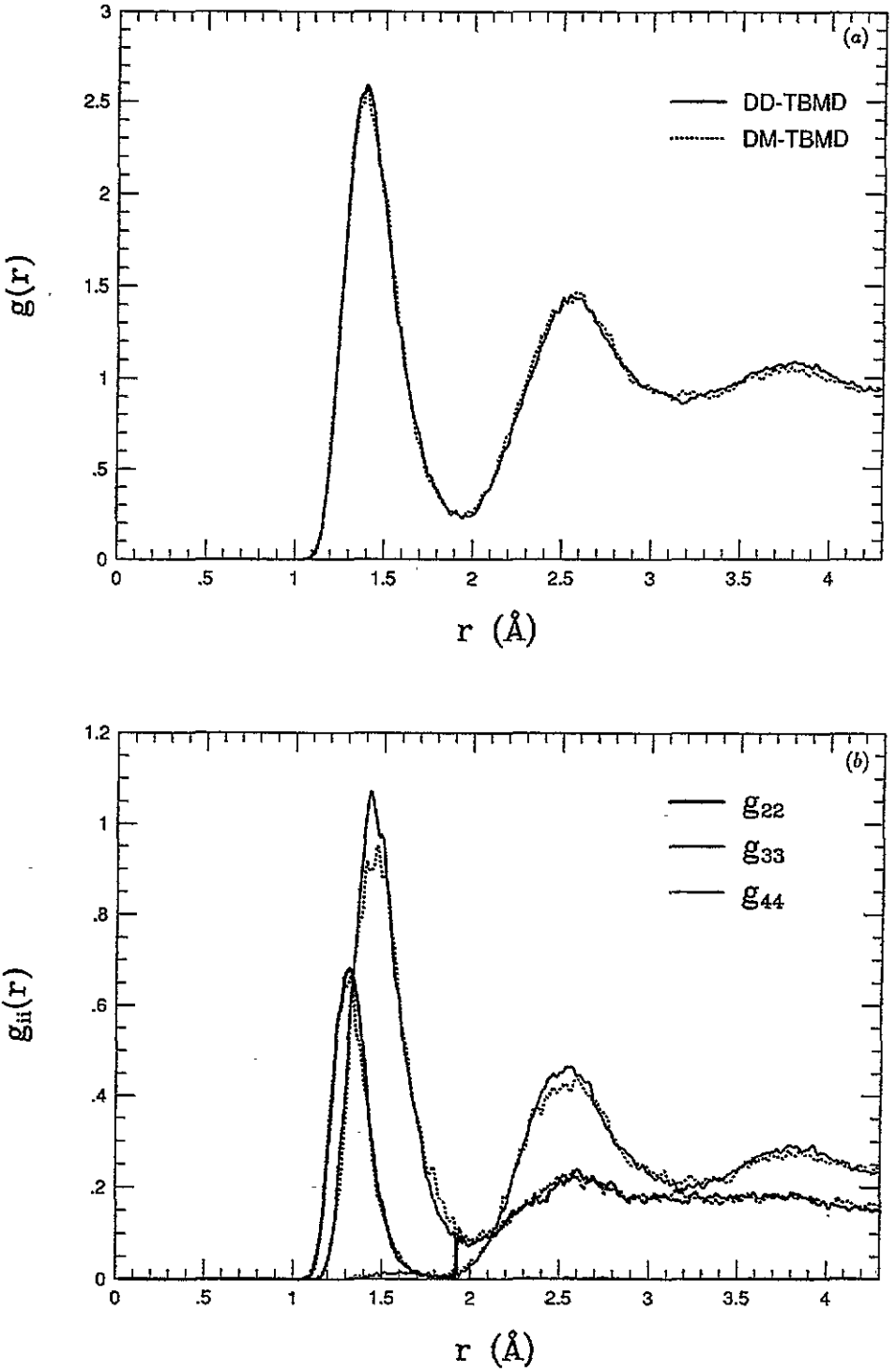

Figure 5. Comparison of the DM-TBMD simulation (broken curves) with the DD-TBMD simulation (full curves) for liquid carbon in (*a*) pair-correlation and (*b*) partial radial distribution.

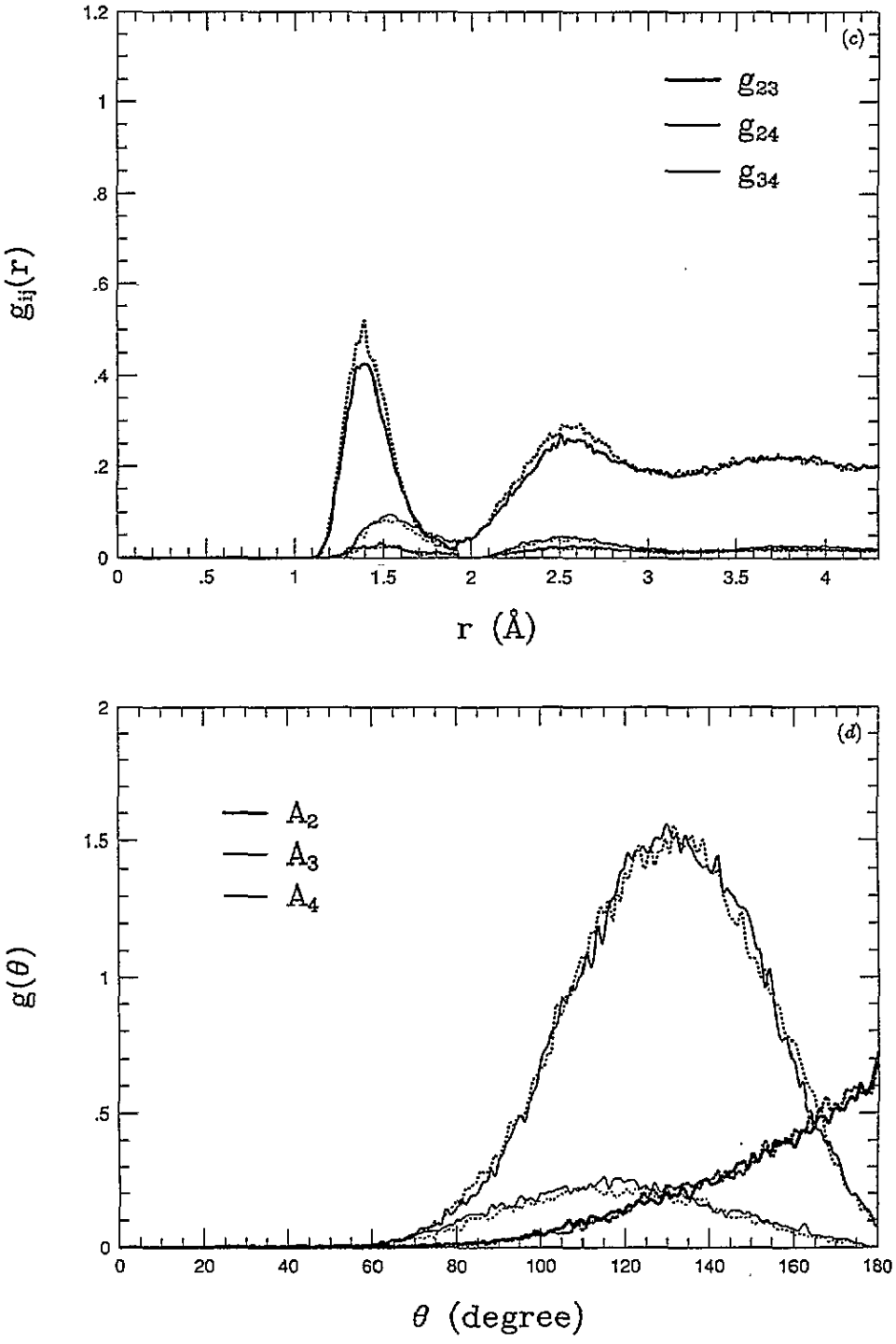

Figure 5. (Continued.) Comparison of the DM-TBMD simulation (broken curves) with the DD-TBMD simulation (full curves) for liquid carbon in (*c*) partial radial distribution and (*d*) angular distribution function.

The smallest possible cutoff range can be estimated by studying the phase diagram of energy against volume curve, or in addition by studying the bulk modulus and optical phonon [1]. For the carbon systems we studied, we chose a cutoff range up to the fifth-neighbour shell of diamond structure, namely $N_c = 46$, at which the DM scheme can reproduce well the energy against volume curve from the DD method. In the MD simulation we found that it is better to keep a constant N_c than to keep a constant R_c in order to reduce an abrupt change of the number of non-zero off-diagonal elements of the density matrix, or equivalently reduce the effect of a large energy discontinuity from one MD timestep to the other.

A good criterion to choose the largest possible tolerance is to check the conservation of the total system energy and other relevant physical properties. It is also related to the choice of the starting variational density matrix for minimization at each MD timestep. The starting density matrix for each MD timestep can always be predicted by the extrapolation technique of (3.1), which can dramatically save the computational effort. However, using such a prediction for the starting density matrix the error tends to be systematic and we found that there is error accumulation from each MD step, which leads to a consistent energy leakage or gain. Although by decreasing the tolerance the degree of leakage or gain of the total energy can be reduced, we found that the non-conservation still persist even up to $\tau = 10^{-10}$. To eliminate the error accumulation in the MD simulation, one can either use the technique of alternating different orders of extrapolation, if they correspond to error accumulation in different directions, or use a 'neutral' scratch density matrix that tends to give random errors and thus does not accumulate errors systematically. For crystalline and amorphous carbon, they correspond to the first case, where the total energy consistently leaks by using the first-order extrapolation and consistently gains by using the second-order extrapolation. Thus, by using these two kinds of extrapolations alternately, we manage to conserve the system energy while using a relatively large tolerance. However, the liquid carbon corresponds to the second case, where we found that we have to use the scratch density matrix as the starting point for minimization at each MD timestep in order to achieve conservation of system energy. For the crystalline and amorphous carbon we studied, by choosing $\tau = 10^{-5}$ and $\tau = 10^{-6}$, we obtained a crossover at around 60 and 120 atoms, respectively. For liquid carbon, by choosing $\tau = 10^{-4}$ but with a scratch density matrix for each MD timestep, we obtained a crossover at around 230 atoms.

The formulation and techniques developed in this paper can also be applied to other schemes with linear system-size scaling, such as the localized-orbital method [3,4]. We have carried out similar calculations using the localized-orbital method and will present the results and comparison with the DM method elsewhere.

As mentioned in section 1, the linear scaling and natural parallelism have allowed the density matrix method to be more efficiently implemented on parallel computers. At the time of submission of this paper, we successfully implemented the DM-TBMD scheme on an Intel Paragon. While optimization is still ongoing, extrapolation from current work indicates that practical simulations of 10 000 atoms systems should be easily achievable on the 512-node XP/S-35, taking an estimated 17 h for 1000 timesteps. As the communication capabilities of the Paragon improve, smaller systems may be spread across all 512 nodes to allow for long simulation times, such as a 1000-atom simulation for 10 000 timesteps taking the same 17 h. This should enable us to perform a wide variety of new problems requiring the study of such a large system and perform long simulations.

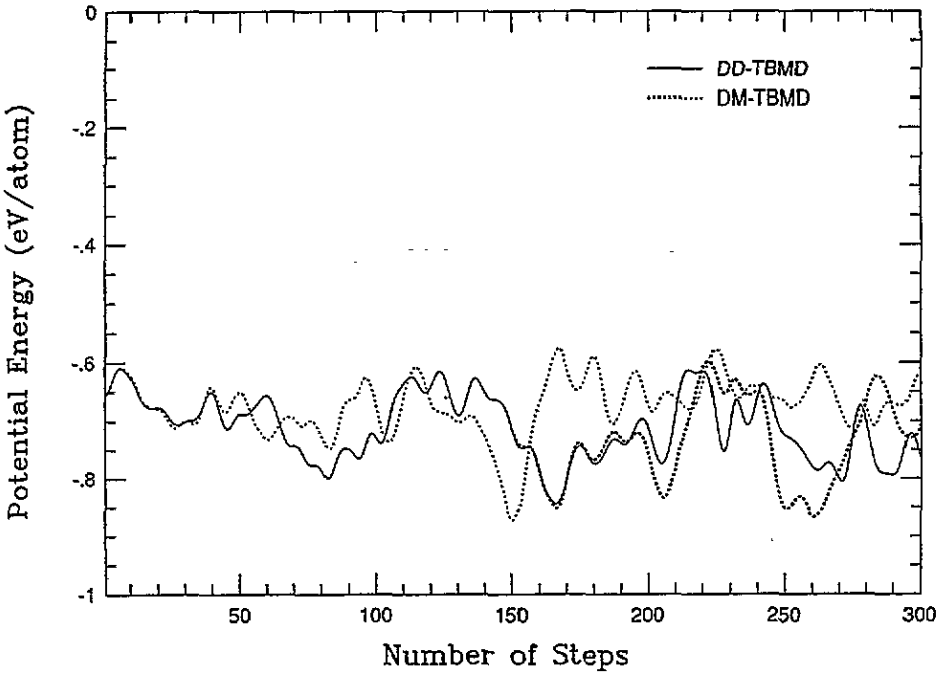

Figure 6. The evolution of potential energy against simulation time for liquid carbon, where the heavy broken curve represents the DM-TBMD result starting from the configuration of DD-TBMD at the 150th timestep.

Acknowledgments

We thank B Harmon, B L Zhang, X D Wang, J Morris and D Turner for helpful discussions and comments. This work is supported by the Director of Energy Research, Office of Basic Energy Sciences, the High Performance Computing and Communications Initiative, and a grant of computation on Cray computers at the NERSC at Livermore. Ames Laboratory is operated for the US Department of Energy by Iowa State University under Contract No W-7405-ENG-82.

Appendix. A generalized density-matrix formulation

In this appendix we derive a generalized density-matrix formulation for electronic-structure calculations for tight-binding models, and show that several O(N) schemes proposed recently can be discussed within a unified framework. Let us consider a tight-binding model with an $N \times N$ Hamiltonian matrix H, and an 'energy' functional defined by

$$\Omega = \mathrm{tr}[f(\bar{\rho})(H - \eta I)] + \eta M \tag{A1}$$

where I is the identity matrix, M is the number of occupied bands in the system ($M < N$) and η is a real number to be specified later. The operator $\bar{\rho}$ is a sum of projectors defined by

$$\bar{\rho} = \sum_{i=1}^{M} |\phi_i\rangle\langle\phi_i| \tag{A2}$$

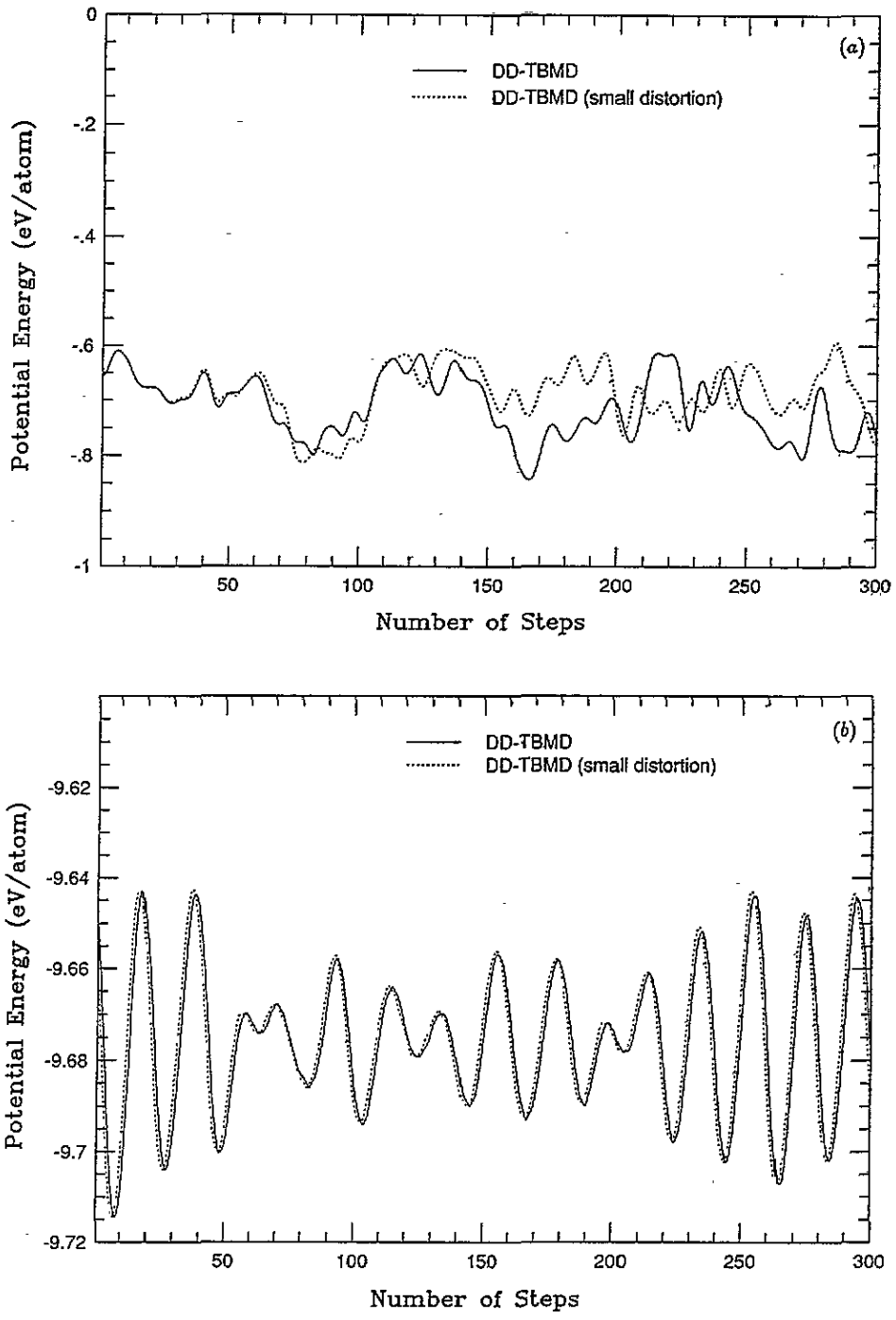

Figure 7. The evolution of potential energy against simulation time for liquid carbon (*a*), and for crystal carbon (*b*), where the broken curves represent the system with slight distortion from the original one indicated by full curves.

where the trial wavefunctions $|\phi_i\rangle$ are linearly independent but not necessarily orthogonal to each other (we will assume that they are normalized since normalization is always a trivial procedure), and $f(\bar{\rho})$ is a functional of the operator $\bar{\rho}$. Thus Ω is an implicit functional of $\{\phi\}$. We will show that for certain proper choices of the function f, $\Omega[\{\phi\}]$ can attain a minimum value of $\sum_{i=1}^{M}\epsilon_i$, where ϵ_i are eigenvalues of the Hamiltonian H, when Ω is minimized by performing an unconstrained minimization with respect to $\{\phi\}$.

Since $\bar{\rho}$ is a Hermitian operator, we can always write

$$\bar{\rho} = \sum_{i=1}^{N} \gamma_i |\psi_i\rangle\langle\psi_i| \tag{A3}$$

where γ_i and $|\psi_i\rangle$ are the eigenvalues and eigenvectors of $\bar{\rho}$, respectively. From (A2), it is easy to see that γ_i are the eigenvalues of the overlap matrix $S_{ij} = \langle\phi_i|\phi_j\rangle$, and hence, $\gamma_i \geqslant 0$. Since $\bar{\rho}$, defined by (A2), only spans an M-dimensional subspace, we may write

$$\bar{\rho} = \sum_{i=1}^{M} \gamma_i |\psi_i\rangle\langle\psi_i| + \sum_{i=M+1}^{N} 0|\psi_i\rangle\langle\psi_i| \tag{A4}$$

that is, it will have $N - M$ eigenvalues with a zero value.

Now for any operator A, we can write

$$\mathrm{tr}[f(\bar{\rho})A] = \sum_{i=1}^{N} f(\gamma_i)\langle\psi_i|A|\psi_i\rangle.$$

Hence, we have

$$\Omega = \mathrm{tr}[f(\bar{\rho})(H - \eta I)] + \eta M \tag{A5}$$

$$= \sum_{i=1}^{N} f(\gamma_i)\big(\langle\psi_i|H|\psi_i\rangle - \eta\big) + \eta M. \tag{A6}$$

Now suppose that we require the function f to satisfy the condition $f(0) = 0$, then

$$\Omega = \sum_{i=1}^{M} f(\gamma_i)\big(\langle\psi_i|H|\psi_i\rangle - \eta\big) + \eta M \tag{A7}$$

$$= \sum_{i=1}^{M}\langle\psi_i|H|\psi_i\rangle + \sum_{i=1}^{M}\big(f(\gamma_i) - 1\big)\big(\langle\psi_i|H|\psi_i\rangle - \eta\big) \tag{A8}$$

$$= \sum_{i=1}^{M} H_{ii} + \sum_{i=1}^{M}\big(f(\gamma_i) - 1\big)\big(H_{ii} - \eta\big). \tag{A9}$$

Since $\sum_{i=1}^{M} H_{ii} \geqslant \sum_{i=1}^{M}\epsilon_i$, where ϵ_i are the eigenvalues of the Hamiltonian H, we have

$$\Omega - \sum_{i=1}^{M}\epsilon_i = P + \sum_{i=1}^{M}\big(f(\gamma_i) - 1\big)(H_{ii} - \eta) \tag{A10}$$

where $P \geqslant 0$ ($P = 0$ when $|\psi_i\rangle$ spans the same subspace as the eigenvectors of H). Now let us choose an $\eta > \epsilon_{\max}$ (where $\epsilon_{\max}$ is the largest eigenvalue of H), then $H_{ii} - \eta = N_i \leqslant 0$ for all i, then

$$\Omega - \sum_{i=1}^{M} \epsilon_i = P + \sum_{i=1}^{M} \big(f(\gamma_i) - 1\big) N_i. \tag{A11}$$

If we choose $f(x)$ such that $f(1)$ is an absolute maximum, then it is quite obvious that Ω will attain a value of $\sum_{i=1}^{M} \epsilon_i$, where it is minimized (without any further constraints) by varying the $\{\phi\}$. The minimization can be performed by steepest-descent or conjugate-gradient algorithm. We note that (A3) is only a schematic representation, thus we do not need to diagonalize the $\bar{\rho}$ matrix at any step. The simplest function satisfying the conditions that $f(0) = 0$ and that $f(1)$ is an absolute maximum, is

$$f(x) = 2x - x^2. \tag{A12}$$

This is basically the formulation used by Mauri and co-workers [3], and Ordejon and co-workers [4], cast in a density matrix formulation for a tight-binding model. There are of course infinitely many possible choices for $f(x)$, with (A12) being the simplest.

So far, we have assumed that we are constructing $\bar{\rho}$ from a set of trial wavefunctions $\{\phi\}$, i.e. $\bar{\rho} = \sum_{i=1}^{M} |\phi_i\rangle\langle\phi_i|$. Now suppose that we do not even specify the trial wavefunction $\{\phi\}$ and use the elements of $\bar{\rho}$ as our basic variables; then we still have (A1) and we can still write $\bar{\rho}$ as in (A3), but γ_i now are not necessarily positive-definite and they may all be non-zero (in the previous case, $\gamma_i = 0$ for $i = M+1, \cdots, N$). Then, we can rewrite (A6) as

$$\Omega = \sum_{i=1}^{M} f(\gamma_i)(H_{ii} - \eta) + \eta M + \sum_{i=M+1}^{N} f(\gamma_i)(H_{ii} - \eta). \tag{A13}$$

Now if we require that (i) η is the chemical potential of the system, (ii) $f(0)$ is the absolute minimum for $f(x)$, (iii) $f(1)$ is the absolute maximum for $f(x)$, (iv) $f(0) = 0$ and (v) $f(1) = 1$, then Ω can achieve a minimum value of $\sum_{i=1}^{M} \epsilon_i$ by varying $\bar{\rho}$. This can be easily seen from (A13) that, at the minimum of Ω, the first term will drive $f(\gamma_i) = f(1) = 1$ since $H_{ii} - \eta < 0$, and the third term will drive $f(\gamma_i) = f(0) = 0$ since $H_{ii} - \eta > 0$. Alternatively, if we change condition (i) to be $\sum_i^M f(\gamma_i) = M$, then Ω can achieve the same minimum value, while η will be driven to the chemical potential of the system. This corresponds to the steepest-descent algorithm described in section 2.

We note that conditions (ii) and (iii) are very difficult to satisfy simultaneously by simple polynomials, so we relax the condition that $f(0)$ and $f(1)$ become the (local) minimum and maximum within a specified range of the eigenvalue spectrum $\{\gamma\}$, and we hope that when we are minimizing Ω by changing the elements of $\bar{\rho}$, the eigenvalues of $\bar{\rho}$ will stay within the specified range. A simple function that satisfies this condition is

$$f(x) = 3x^2 - 2x^3. \tag{A14}$$

This function, as also shown in (2.3), is the essence of the density matrix method proposed by Li and co-workers [1] and Daw [2]. In this case, $\sum_{i=1}^{M} \epsilon_i$ is just a local minimum of Ω. The absolute minimum is actually negative infinity, and hence there is always the possibility of 'runaway' solutions.

The method proposed by Wang and Teter [5] is also very similar to what we have discussed above. They propose to minimize

$$\Omega = \sum_{i=1}^{M} \langle \phi_i | H | \phi_i \rangle + \eta \sum_{i,j \neq i}^{M} (\langle \phi_i | \phi_j \rangle)^2. \tag{A15}$$

This is equivalent to

$$\Omega = \mathrm{tr}[\bar{\rho} H] + \eta \mathrm{tr}[\bar{\rho}^2 - \bar{\rho}] \tag{A16}$$

where $\bar{\rho} = \sum_{i=1}^{M} |\phi_i\rangle\langle\phi_i|$, while (A1) with the choice of $f(x) = 2x - x^2$ becomes

$$\Omega = \mathrm{tr}[\bar{\rho}(H - \eta I)] + \eta M + \eta \mathrm{tr}[\bar{\rho}^2 - \bar{\rho}] + \mathrm{tr}[(\bar{\rho} - \bar{\rho}^2) H]. \tag{A17}$$

Since at the minimum of the above equation, $\mathrm{tr}(\bar{\rho}) = M$, basically the formulation of Wang–Teter missed the fourth term of (A17). We note that (A17), as a special case of (A1) with a certain 'proper' choice of $f(x)$, will give us the 'exact' result of $\Omega = \sum_{i=1}^{M} \epsilon_i$ for a tight-binding model, but the method of Wang–Teter (equations (A15) and (A16)) will not. When we minimize the Wang–Teter functional with respect to $\{\phi\}$, the functional will try to find a compromise between minimizing the band energy $\mathrm{tr}(\bar{\rho} H)$ and enforcing idempotency $\bar{\rho}^2 = \bar{\rho}$, and the final result in fact will depend on the choice of η.

In our discussions we do not limit the range of $\bar{\rho}$ (or $\{\phi\}$) in real space. In practice, such truncation is essential to make these algorithms scale linearly with system size. There is of course always a truncation error, which needs to be balanced with the efficiency of the method.

References

[1] Li X P, Nunes R W and Vanderbilt D 1993 *Phys. Rev.* B **47** 10 891
[2] Daw M S 1993 *Phys. Rev.* B **47** 10 895
[3] Mauri F, Galli G and Car R 1993 *Phys. Rev.* B **47** 9973
[4] Ordejón P, Drabold D, Grumbach M P and Martin R M 1993 *Phys. Rev.* B **48** 14 646
[5] Wang L W and Teter M P 1992 *Phys. Rev.* B **46** 12 798
[6] Baroni S and Giannozzi P 1992 *Europhys. Lett.* **17** 547
[7] Yang W 1991 *Phys. Rev. Lett.* **66** 1438
[8] Baroni S and Giannozzi P 1992 *Europhys. Lett.* **17** 547
[9] Goedecker S and Colombo L 1994 *Phys. Lett.* **73** 122
[10] Wang C Z, Chan C T and Ho K M 1989 *Phys. Rev.* B **39** 8586
[11] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[12] Xu C H, Wang C Z, Chan C T and Ho K M 1992 *J. Phys.: Condens. Matter* **4** 6047
[13] Khan F S and Broughton J Q 1989 *Phys. Rev.* B **39** 3688
[14] Chadi D J 1984 *Phys. Rev.* B **29** 785
[15] Chadi D J and Martin R M 1976 *Solid State Commun.* **19** 643
[16] Wang C Z, Ho K M and Chan C T 1993 *Phys. Rev.* B **47** 14 835
[17] Zhang B L, Wang C Z, Ho K M, Xu C H and Chan C T 1992 *Chem. Phys. Lett.* **97** 5007
[18] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045
[19] Arias T A, Payne M C and Joannopoulos J D 1992 *Phys. Rev.* B **45** 1538
[20] Wang C Z and Ho K M 1993 *Phys. Rev. Lett.* **71** 1184