Large-scale efficient Langevin dynamics, and why it works

JIAYU DAI and JIANMIN YUAN[a]

Department of Physics, National University of Defense Technology - Changsha 410073, PRC

Abstract – An effective method with large convergence tolerance for self-consistent-field calculations is proposed to accelerate Born-Oppenheimer molecular-dynamics simulations based on the density functional theory. The errors for forces with large convergence tolerance are taken as the noises of Langevin dynamics for constant-temperature molecular dynamics. We use different systems such as crystalline Si, liquid Si and liquid water to verify the accuracy and efficiency of the method. Results of the new method are consistent with those of sufficiently accurate methods and the dynamical properties of the systems are reproduced as well. The efficiency of computation is improved by a factor of two to three using this model. Our analysis provides three reasons which explain why the method works. This development effectively extends the application of Born-Oppenheimer molecular dynamics to larger systems.

Ab initio molecular dynamics based on density functional theory (DFT) [1] is a feasible approach in a variety of studies in physics, geology, material science, chemistry and biology. However, its computational cost is very expensive, especially for Born-Oppenheimer molecular dynamics (BOMD) [2], which limits its applications in both time scale and dimension scale. In order to avoid fully minimizing the DFT functional in every molecular-dynamics time step, Car-Parrinello molecular dynamics (CPMD) [3] were formulated. However, the time step in the CPMD scheme must be much shorter than that used in BOMD. In BOMD, the motions of atoms are driven by forces that are calculated in the self-consistent electronic ground state. In order to guarantee the ground state for each instantaneous arrangement of ions, highly accurate convergence with many iteration steps is required in the self-consistent-field (SCF) calculation, which is one of the primary reasons for the expensive computational cost in BOMD. If the convergence of the SCF calculation is not accurate enough, the true dynamical process of the system would be destroyed. Thus, how to retain the correct dynamical properties with a large convergence tolerance and fewer iteration steps than in BOMD becomes a key problem in reducing the computational cost. Some work for accelerating BOMD has been done during recent years [4–11]. Florian *et al.* introduced noise forces to the linear scaling electronic structure

calculations [12], which acquired dynamical properties without exact forces. Kühne *et al.* developed an efficient and accurate method combining the accuracy and large time step of BOMD, resulting in an efficiency comparable to CPMD [11] using this idea. Nevertheless, these authors did not explain why these modified methods worked well.

In this paper, a method with large convergence tolerance of the SCF calculation is presented for performing Born-Oppenheimer molecular dynamics. An explanation as to why this method works is given, which is important for the comprehension of the method. The results of all tests performed using this method are consistent with those of accurate convergence. Furthermore, the dynamical properties are reproduced satisfactorily, and the efficiency is improved by a factor of two to three as well.

It is widely acknowledged that if a set of good initial wave functions is used at the beginning of each SCF calculation, the exact Born-Oppenheimer potential energy surface can be obtained quickly. Extrapolation can solve this problem partly by guessing the initial functions [4]. The wave functions $\boldsymbol{\psi}(t + \mathrm{d}t)$ at the next time step can be estimated from the previous steps:

$$\boldsymbol{\psi}(t + \mathrm{d}t) = \boldsymbol{\psi}(t) + \alpha(\boldsymbol{\psi}(t) - \boldsymbol{\psi}(t - \mathrm{d}t))$$
$$+ \beta(\boldsymbol{\psi}(t - \mathrm{d}t) - \boldsymbol{\psi}(t - 2\mathrm{d}t)), \qquad (1)$$

where t is the present time, $\mathrm{d}t$ is the molecular-dynamics (MD) integration step and α and β are coefficients defined previously in ref. [4].
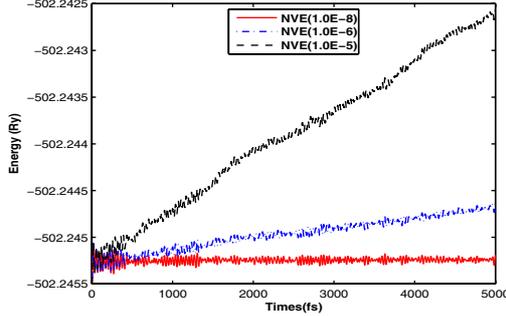
---

[a]E-mail: jmyuan@nudt.edu.cn

Fig. 1: (Color online) The energy drift of an Si system with 64 atoms using NVE calculations with convergences of 10^{-8}, 10^{-6} and 10^{-5}, respectively.

The initial electronic charge density $\boldsymbol{\rho}(t+\mathrm{d}t)$ can be extrapolated as well. According to Alfè's method [6], the charge density can be estimated using

$$\boldsymbol{\rho}(t+\mathrm{d}t) = \boldsymbol{\rho}_{at}(t+\mathrm{d}t) + \delta\boldsymbol{\rho}(t+\mathrm{d}t), \qquad (2)$$

where $\boldsymbol{\rho}(t+\mathrm{d}t)$ is the self-consistent charge density at time $t+\mathrm{d}t$, $\boldsymbol{\rho}_{at}(t+\mathrm{d}t)$ is the sum of the atomic charge densities at time $t+\mathrm{d}t$ and $\delta\boldsymbol{\rho}(t+\mathrm{d}t)$ is given by

$$\delta\boldsymbol{\rho}(t+\mathrm{d}t) = \delta\boldsymbol{\rho}(t) + \alpha(\delta\boldsymbol{\rho}(t) - \delta\boldsymbol{\rho}(t-\mathrm{d}t))$$
$$+ \beta(\delta\boldsymbol{\rho}(t-\mathrm{d}t) - \delta\boldsymbol{\rho}(t-2\mathrm{d}t)), \qquad (3)$$

where α and β are the same coefficients as in eq. (1). The potential energy surface on which the ions move is defined by the minimum of the energy functional $\mathbf{E}_{DFT}[\boldsymbol{\psi},\boldsymbol{\rho}]$. According to DFT, both $\boldsymbol{\psi}$ and $\boldsymbol{\rho}$ are functions of the ionic positions $\mathbf{R}_I(t)$. In this notation, the equations of motion in the Born-Oppenheimer molecular dynamics are

$$M_I\ddot{\mathbf{R}}_I = -\nabla_I \min \mathbf{E}_{DFT}[\boldsymbol{\psi},\boldsymbol{\rho}], \qquad (4)$$

where M_I is the I-th ionic mass.

The efficiency of minimizing the energy $\mathbf{E}_{DFT}$ can be improved significantly using this extrapolation. Yet, since the tolerance of convergence has to be sufficiently small, the number of SCF iterations is generally still too large. If a larger tolerance of convergence can be used, the efficiency could be further improved. If this is adopted, the energy and forces of the system calculated with a larger convergence tolerance would be drifted or noisy. The drift of the total energy (kinetic energy plus potential energy) in the NVE ensemble with different convergence is shown in fig. 1, where the drift with large tolerance is apparent.

Similar consequences result for the forces. The characteristics of forces on atoms can be understood from two aspects: 1) What will happen when the tolerance is not small enough? In order to understand this question, the distribution of the force errors F (the difference between the forces obtained with insufficient convergence (10^{-4}) and sufficient convergence (10^{-8}) in an identical SCF calculation) at every time step within 8000 fs is plotted in the upper panel of fig. 2, which shows typical
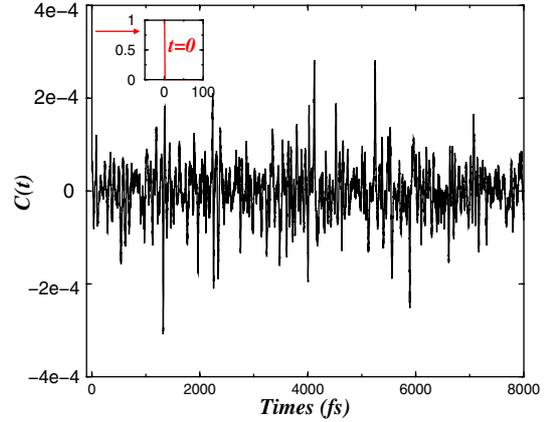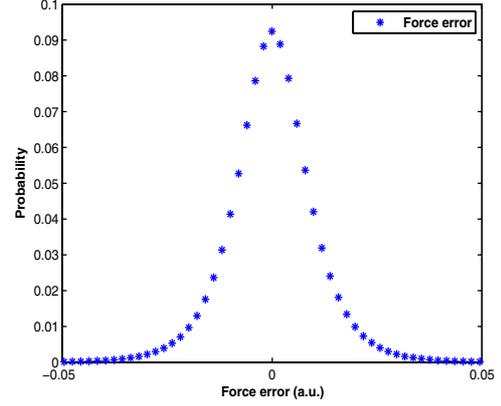

Fig. 2: (Color online) The upper graph shows the distribution of force errors of an Si system with 64 atoms using NVT calculations (and a velocity re-scaling method) between small (10^{-8}) and large (10^{-4}) convergence tolerances at 3000 K; the lower graph shows the normalized autocorrelation function of the errors in the force with an inner plot showing the slice at $t=0$.

Gaussian distribution. Then, the autocorrelation function of the force errors, $\mathbf{C}(t) = \langle\mathbf{F}(t+\tau)\cdot\mathbf{F}(\tau)\rangle$, is calculated and shown in the lower panel of fig. 2. It can be seen that the normalized autocorrelation is very small (around 10^{-4}) if $t\neq0$ and it has typical white noise features, i.e., $\mathbf{C}(0) \gg \mathbf{C}(t\neq0)$. Therefore, it can be deduced that the force errors can be considered as a Gaussian white noise. 2) What is the correlation between the force errors and the velocities of ions? The value of the correlation coefficient between the force errors and the velocities of all the atoms, i.e., $\langle\mathbf{F}\cdot\mathbf{V}\rangle$, is equal to 1.2×10^{-9}. From this value we conclude that they are uncorrelated with each other. For water at 325 K (32 water molecules), we reach the same conclusion. Due to these two reasons, the force errors can be sampled directly as white noise with a Gaussian distribution, which will result in the exact Boltzmann distribution below.

Thus, these force errors can be corrected if we assume that the forces from the dynamics $\mathbf{F}_{TR}$ can be taken as $\mathbf{F}_{TR} = \mathbf{F}_{BO} + \mathbf{N}_I^D$, where $\mathbf{F}_{BO}$ is the force obtained

from the SCF calculation with sufficient convergence, and $\mathbf{N}_I^D$ is random noise coming from the damping forces when the convergence is insufficient for BOMD. According to the analysis of force errors above, it can be considered as a white noise. That is why we can treat the force errors as thermostat in the constant-temperature molecular dynamics. Hence, one can perform molecular dynamics with the canonical ensemble using the Langevin equation:

$$M_I\ddot{\mathbf{R}}_I = \mathbf{F}_{BO} - \gamma_L\dot{\mathbf{R}}_I + \mathbf{N}_I^L, \qquad (5)$$

where γ_L is a Langevin friction coefficient, and $\mathbf{N}_I^L$ is a Gaussian random noise. This equation can be written as

$$M_I\ddot{\mathbf{R}}_I = \mathbf{F}_{BO} - (\gamma_L + \gamma_D)\dot{\mathbf{R}}_I + \mathbf{N}_I, \qquad (6)$$

where $\mathbf{N}_I = \mathbf{N}_I^D + \mathbf{N}_I^L$ is a Gaussian random noise, and γ_D is the intrinsic friction coefficient corresponding to the random noise $\mathbf{N}_I^D$.

To be assured of an accurate sampling of the Maxwell-Boltzmann distribution, the noise must obey the fluctuation dissipation theorem:

$$\langle\mathbf{N}_I(0)\mathbf{N}_I(t)\rangle = 6(\gamma_D + \gamma_L)M_I k_B T \mathrm{d}t, \qquad (7)$$

where γ_L is arbitrary, and the coefficient γ_D can be determined by varying it until the equipartition theorem $\langle\frac{1}{2}M_I\dot{\mathbf{R}}_I^2\rangle = \frac{3}{2}k_B T$ is satisfied. Here, the choice of γ_D has to guarantee that the noises are smaller than the forces on the atoms, in order to preserve the dynamical behavior of the system. Once γ_D is determined, it has to be kept constant during the calculation. With this choice, the drifted forces with larger tolerance of convergence will be corrected and the force errors can be used in the sampling. Through this procedure, one can exactly calculate static observables within the framework of Langevin dynamics without knowing the exact forces. Thanks to this treatment, molecular-dynamics simulations can be done by performing SCF calculations at each time step with a large tolerance of convergence and fewer iterations.

In order to integrate eq. (6), we use the algorithm in a Verlet-like form [13]:

$$\mathbf{R}_I(t+\mathrm{d}t) = \mathbf{R}_I(t) + (\mathbf{R}_I(t) - \mathbf{R}_I(t-\mathrm{d}t))\frac{1 - \frac{1}{2}\gamma_T\mathrm{d}t}{1 + \frac{1}{2}\gamma_T\mathrm{d}t}$$

$$+ (\mathrm{d}t^2/M_I)(\mathbf{F}_{BO}(t) + \mathbf{N}_I(t)\left(1 + \frac{1}{2}\gamma_T\mathrm{d}t\right)^{-1}, \qquad (8)$$

where t and $\mathrm{d}t$ are the same as the eq. (1), and $\gamma_T = \gamma_L + \gamma_D$. The velocities $\mathbf{v}_I(t+\mathrm{d}t)$ of the ions can be calculated by using the Verlet formula

$$\mathbf{v}_I(t+\mathrm{d}t) = \dot{\mathbf{R}}_I = \frac{\mathbf{R}_I(t+\mathrm{d}t) - \mathbf{R}_I(t-\mathrm{d}t)}{2\mathrm{d}t}. \qquad (9)$$

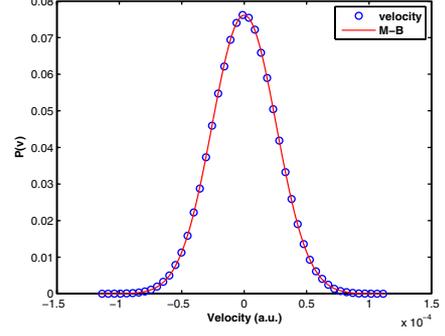This method has been implemented in the plane waves code PWSCF [14]. In our calculations, we set $\gamma_L = 0$

Fig. 3: (Color online) The velocity distribution of an Si system with 64 atoms at $300\,\mathrm{K}$. M-B is the Maxwell-Boltzmann distribution, and the circles represent the calculated velocity distribution.

for all time, and the intrinsic friction coefficient γ_D is chosen to be small enough (about $10^{-8}\,\mathrm{fs}^{-1}$) to keep the random forces smaller than the systematic forces. Since it is so small, we can reproduce the dynamical procedure smoothly and correctly, and keep the influence of thermostat small as well.

To verify our approach, systems of 64 atoms of liquid Si ($3000\,\mathrm{K}$), crystalline Si ($300\,\mathrm{K}$) and 32 water molecules ($325\,\mathrm{K}$) were chosen, respectively, for simulation. All simulations have been performed using the basis of the plane wave expansion, adequate energy cutoff, and norm-conserving pseudopotentials with Perdew-Burke-Ernzerhof (PBE) exchange correlation functional [15] for Si, and ultrasoft pseudopotentials [16] with Perdew-Zunger (PZ) [17] exchange correlation functional for H and O. Equation (6) was integrated via eq. (8) with an integration MD time step of $\mathrm{d}t = 1\,\mathrm{fs}$. For simplicity, the Brillouin zone was sampled at the Γ-point only. The intrinsic friction coefficient is $\gamma_D \sim 10^{-8}\,\mathrm{fs}^{-1}$ at the temperatures $300\,\mathrm{K}$ and $325\,\mathrm{K}$, and $\gamma_D \sim 10^{-6}\,\mathrm{fs}^{-1}$ at $3000\,\mathrm{K}$.

The sampling of the canonical ensemble is considered first. The velocity distribution of the 64 atom Si system at $300\,\mathrm{K}$ for a trajectory of time length $10000\,\mathrm{fs}$ is shown in fig. 3. The velocity distribution is in good agreement with the exact Maxwell-Boltzmann distribution; that is to say, we are indeed performing a canonical sampling, and the damped forces can be taken as thermostat in the NVT ensemble. This can also be concluded from the white noise of the force errors directly, whose properties are the same as the intrinsic noise in the Langevin equation. It is worth pointing out that from the Boltzmann distribution only, one can not definitely deduce that the distribution of force errors is the white noise. There are two possible distributions for the force errors: 1) white noise, and 2) the force errors are linearly proportional to the velocity, that is, $\mathbf{F}_{TR} = \mathbf{F}_{BO} - \gamma_D\dot{\mathbf{R}}_I$, both of which result in the correct Boltzmann average in the Langevin dynamics. Therefore, our analysis of the Gaussian distribution and the uncorrelated relation between force errors and velocities
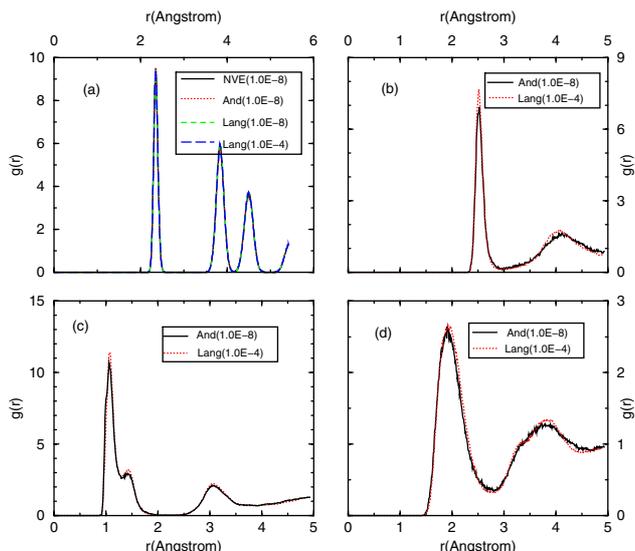
Fig. 4: (Color online) The radial distribution function (RDF). "NVE(1.0E-8)" refers to the NVE results using $\gamma_D = 0$ (free dynamics) with tolerance of 10^{-8}, "And(1.0E-8)" to the results using Andersen thermostat with tolerance of 10^{-8}, "Lang(1.0E-8)" refers to the results using Langevin dynamics with tolerance of 10^{-8} and "Lang(1.0E-4)" refers to the results using Langevin dynamics with tolerance of 10^{-4}. (a) The RDF of a 64 atom crystalline Si system at temperature 300 K; (b) the RDF of H-H partial of the 32 molecule water system at temperature 325 K; (c) the RDF of H-O partial of the 32 molecule water system at temperature 325 K; (d) the RDF of O-O partial of the 32 molecule water system at a temperature of 325 K.

ascertains the validity of modeling force errors as white noise and explains the origin of the exact Boltzmann average during the process of molecular-dynamics simulation.

The correct distribution of velocities shows that our model can be used as a thermostat when the convergence is accurate enough. Below, a comparison of results between accurate convergence and insufficient convergence tells us that large tolerance of convergence is feasible in the simulations.

Firstly, the radial distribution functions (RDF) with different SCF convergence tolerances were calculated and are shown in fig. 4. In these calculations, Andersen thermostat [18] and NVE ensemble are used to do the comparison. Compared with that of small tolerance (10^{-8}), the RDF of large tolerance (10^{-4}) is adequate and the use of this tolerance does not degrade the quality of the simulation. The friction coefficient was set to be $\gamma_D \sim 10^{-8}\,\mathrm{fs}^{-1}$.

Secondly, some dynamical trajectories of the systems were analyzed based on the calculations. The temperature and the total energy (kinetic energy plus potential energy) trajectories are given in fig. 5 and fig. 6, respectively, which show that these properties are reproduced well by the present model. In fig. 5, the instantaneous temperature fluctuates around the target temperature (300 K), indicating that the correct thermostat was performed. Considering the Si system at temperature of 3000 K as a
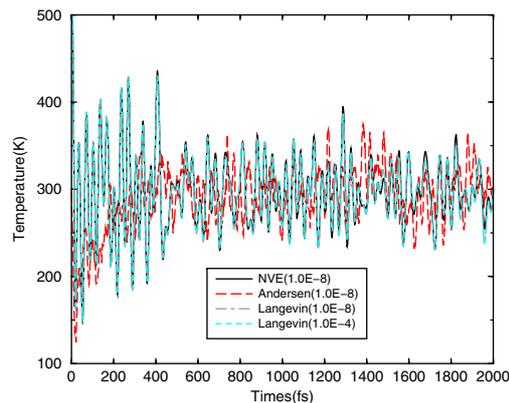

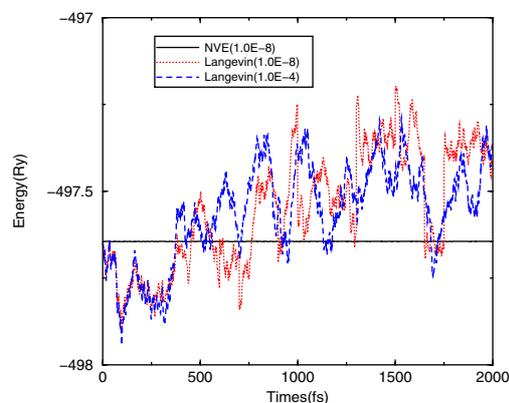Fig. 5: (Color online) The temperature trajectory of a 64 atom Si system at a temperature of 300 K.

Fig. 6: (Color online) The total energy (kinetic energy plus potential energy) trajectory for a liquid Si system with 64 atoms at a temperature of 3000 K, calculated with Langevin friction coefficient $\gamma_L = 0$ and intrinsic friction coefficient $\gamma_D \sim 10^{-6}\,\mathrm{fs}^{-1}$.

metallic system, we used 30 Ry as the kinetic energy cutoff. It is seen from fig. 6 that the energy with thermostat is close to the energy of the NVE calculation. Furthermore, the energy with Langevin thermostat of large tolerance (10^{-4}) is close to the energy with low tolerance (10^{-8}). This performance ensures that the potential energy surface can be kept close to the Born-Oppenheimer potential energy surface and guarantees all properties of the system are satisfactory even though SCF calculations with large tolerance of convergence are involved in the present approach.

Finally, the computational cost can be reduced considerably by using this model. The number of SCF iterations in different MD time steps at different temperatures can be reduced. The number of SCF iterations in our model (Langevin with tolerance of 10^{-4}) is more than two times fewer than that of others as shown in fig. 7. For the system of 32 water molecules, the number of iterations in each time step was almost three, whereas more than six were required in the other methods. The computational cost is reduced by a factor of two to three here. In
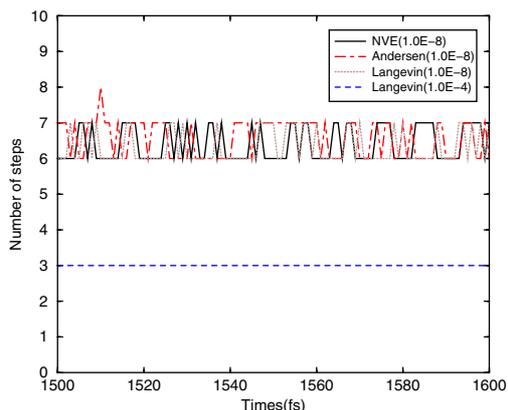
Fig. 7: (Color online) The number of SCF steps in each MD time step for a liquid water system with 32 molecules at a temperature of 325 K.

fact, the speed-up of this method can be increased by improving the extrapolation method, such as the Always Stable Predictor-Corrector (ASPC) extrapolation [19], or improving the time reversible properties [10].

In conclusion, an effective method using large SCF convergence tolerance of the electronic structure in the Born-Oppenheimer molecular dynamics has been presented. According to this model, all results are in good agreement with those using the usual accurate convergence with computational efficiency accelerated by a factor of two to three. The number of electronic SCF iterations for each MD time step is reduced by the combination of optimizing the wave functions and using a large tolerance of convergence. The analysis showed that the method works well because of three reasons: 1) The force errors are Gaussian white noises; 2) An uncorrelated relation exists between the force errors and the velocity, and 3) The total energy of the system is kept close to that of the true ground state. The first two reasons explain why we can directly sample the force errors to be white noise, which results in the correct Boltzmann distribution of velocity in the MD simulation. The third reason indicates why this method has useful and satisfactory properties. Using the method presented here, the application of *ab initio* molecular dynamics can be extended to larger systems. Recently, we realized that colored noise can improve the efficiency of CPMD [9], with hints at possible application in BOMD.

REFERENCES

[1] KOHN W. and SHAM L. J., *Phys. Rev.*, **140** (1965) A1133.
[2] MARX D. and HUTTER J., *Modern Methods and Algorithms of Quantum Chemistry*, edited by GROTENDORST J. (J. von Neumann Institute for Computing, Jülich, Germany) 2000.
[3] CAR R. and PARRINELLO M., *Phys. Rev. Lett.*, **55** (1985) 2471.
[4] ARIAS T. A., PAYNE M. C. and JOANNOPOULOS J. D., *Phys. Rev. Lett.*, **69** (1992) 1077; *Phys. Rev. B*, **45** (1992) 1538.
[5] CHAN C. T., BOHNEN K. P. and HO K. M., *Phys. Rev. B*, **47** (1993) 4771.
[6] ALFÈ D., *Comput. Phys. Commun.*, **118** (1999) 31.
[7] VANDEVONDELE J., KRACK M., MOHAMED F., PARRINELLO M., CHASSAING T. and HUTTER J., *Comput. Phys. Commun.*, **167** (2005) 103.
[8] PULAY P. and FOGARASI G., *Chem. Phys. Lett.*, **386** (2004) 272.
[9] CERIOTTI M., BUSSI G. and PARRINELLO M., *Phys. Rev. Lett.*, **102** (2009) 020601.
[10] NIKLASSON A. M. N., TYMCZAK C. J. and CHALLACOMBE M., *Phys. Rev. Lett.*, **97** (2006) 123001.
[11] KÜHNE T. D., KRACK M., MOHAMED F. R. and PARRINELLO M., *Phys. Rev. Lett.*, **98** (2007) 066401.
[12] KRAJEWSKI F. R. and PARRINELLO M., *Phys. Rev. B*, **73** (2006) 041105(R).
[13] PASTOR R. W., BROOKS B. R. and SZABO A., *Mol. Phys.*, **65** (1988) 1409.
[14] GIANNOZZI P. *et al.*, *J. Phys.: Condens. Matter*, **21** (2009) 395502.
[15] PERDEW J. P., BURKE K. and ERNZERHOF M., *Phys. Rev. Lett.*, **77** (1996) 3865.
[16] VANDERBILT D., *Phys. Rev. B*, **41** (1990) 7892.
[17] PERDEW J. P. and ZUNGER A., *Phys. Rev. B*, **23** (1981) 5048.
[18] FRENKEL D. and SMIT B., *Understanding Molecular Simulation* (Acedemic Press, San Diego, USA) 1996.
[19] KOLAFA J., *J. Comput. Chem.*, **25** (2004) 335.