Divide-and-conquer density functional theory on hierarchical real-space grids: Parallel implementation and applications

Fuyuki Shimojo,[1,2] Rajiv K. Kalia,[1] Aiichiro Nakano,[1] and Priya Vashishta[1]
[1]*Collaboratory for Advanced Computing and Simulations, Department of Physics and Astronomy, Department of Computer Science, and Department of Materials Science and Engineering, University of Southern California, Los Angeles, California 90089-0242, USA*
[2]*Department of Physics, Kumamoto University, Kumamoto 860-8555, Japan*
(Received 18 September 2007; published 5 February 2008)

A linear-scaling algorithm based on a divide-and-conquer (DC) scheme has been designed to perform large-scale molecular-dynamics (MD) simulations, in which interatomic forces are computed quantum mechanically in the framework of the density functional theory (DFT). Electronic wave functions are represented on a real-space grid, which is augmented with a coarse multigrid to accelerate the convergence of iterative solutions and with adaptive fine grids around atoms to accurately calculate ionic pseudopotentials. Spatial decomposition is employed to implement the hierarchical-grid DC-DFT algorithm on massively parallel computers. The largest benchmark tests include 11.8×10^6-atom (1.04×10^{12} electronic degrees of freedom) calculation on 131 072 IBM BlueGene/L processors. The DC-DFT algorithm has well-defined parameters to control the data locality, with which the solutions converge rapidly. Also, the total energy is well conserved during the MD simulation. We perform first-principles MD simulations based on the DC-DFT algorithm, in which large system sizes bring in excellent agreement with x-ray scattering measurements for the pair-distribution function of liquid Rb and allow the description of low-frequency vibrational modes of graphene. The band gap of a CdSe nanorod calculated by the DC-DFT algorithm agrees well with the available conventional DFT results. With the DC-DFT algorithm, the band gap is calculated for larger system sizes until the result reaches the asymptotic value.

DOI: 10.1103/PhysRevB.77.085103 PACS number(s): 71.15.−m, 02.70.−c, 02.70.Ns

I. INTRODUCTION

Mechanically induced chemical reactions are important in many material processes.[1] However, theoretical study of mechanochemical processes is challenging because of the interplay between chemical reactions, which are local in nature, and long-range stress phenomena. Examples are energetic materials, in which chemical reactions sustain shock waves,[2–4] and stress corrosion cracking, where chemical reactions at the crack tip are inseparable from long-range stress fields.[5] Furthermore, these chemical reactions are sensitive to material microstructures such as shock-front structures, voids, and grain boundaries.[4] These factors require molecular-dynamics (MD) simulations involving multimillion atoms,[6–8] in which chemical reactions must be described quantum mechanically.[9,10]

One computational approach to quantum mechanically informed multimillion-atom MD simulations[4,11] employs environment-dependent interatomic potentials[12] based on reactive bond orders[13,14] for chemical bond formation and breakage as well as variable atomic charges[15–19] to describe charge transfers. In the reactive force field (ReaxFF) approach,[12] the parameters in the interatomic potential are "trained" to best fit thousands of quantum-mechanical (QM) calculations on small ($N\sim10$) clusters of various atomic-species combinations.[3] Thus, the determination of the potential parameters that describe local chemistry in addition to the bulk mechanical behavior of the material prior to MD simulations constitutes the most critical and nontrivial part of the ReaxFF MD approach. In addition, results of these semiempirical MD simulations need to be validated against more accurate QM calculations, e.g., those based on the density functional theory (DFT).[20–27]

Alternatively, DFT calculations can be performed "on the fly" to compute interatomic forces quantum mechanically during a MD simulation.[28–30] Unfortunately, DFT-based MD simulations are rarely performed for $N>10^2$ atoms because of the $O(N^3)$ computational complexity, which severely limits their applicability. To reduce such high computational complexity, we have recently developed a divide-and-conquer (DC) on cellular decomposition framework to design linear-scaling algorithms (in which the computation time is proportional to the problem size) on massively parallel computers for a broad range of applications.[11] In the context of the DFT, it amounts to the DC-DFT algorithm,[31–33] which represents the physical system as a union of overlapping spatial domains and computes physical properties as linear combinations of domain properties. The DC-DFT algorithm is based on a data locality principle called the quantum near-sightedness,[34] which naturally leads to $O(N)$ QM calculations.[35–59] In this paper, we present a hierarchical-grid DC-DFT algorithm,[60] in which DFT calculations of atomic clusters (or domains) involved in the DC scheme are performed using a real-space approach that numerically represents electronic wave functions on grid points.[61,62] The real-space grid is augmented with a coarser multigrid to accelerate the convergence of iterative solutions.[10,30,38,63–65] Furthermore, finer grids are adaptively generated near atoms[5,66] in order to accurately operate ionic pseudopotentials[67–69] for describing electron-ion interactions. We include electron-electron interactions using the generalized gradient approximation[70] to the exchange-correlation energy. Since the DC-DFT algorithm involves solutions to small (i.e., the average number of atoms per domain, $n<50$) electronic-structure problems, it is free from some of

the convergence problems often associated with iterative solutions in large electronic-structure calculations. For example, it is difficult to achieve convergence for nontrivial electronic-structure problems such as those in amorphous materials with a large number ($>10^4$) of atoms using a large (10^5 grid points per electronic wave function) basis set, unless a good initial guess for the wave functions and the electron number density is available.[29] Our numerical tests show that the DC-DFT algorithm has robust convergence properties such that, starting from random initial wave functions, converged solutions are obtained for such problems. The DC-DFT algorithm has a well-defined set of localization parameters (the size of a domain and the width of buffer layers to augment each domain for avoiding artificial boundary effects), with which the energy converges rapidly. The DC-DFT MD algorithm does not suffer from the energy drift problem, which plagues many $O(N)$ DFT-based MD algorithms, especially with large basis sets ($>10^5$ unknowns per electron, necessary for the transferability of accuracy).

The DC-DFT algorithm on the hierarchical real-space grids is implemented on massively parallel computers using spatial decomposition,[8,11,38] in which the physical system is divided into subsystems of equal volume and each subsystem is assigned to a processor or a small group of processors in a parallel computer. Each subsystem (or processor) contains one or more domains of the DC-DFT algorithm. For each domain, an $O(n^3)$ DFT algorithm is employed to calculate its electronic structures, with little information needed from other processors. The resulting large computation/communication ratio makes this approach highly scalable on massively parallel computers.

The robust convergence properties of the hierarchical-grid DC-DFT algorithm and its excellent scalability on parallel computers have allowed us to perform first-principles MD simulations of several large systems, including liquid rubidium (Rb), graphene, and cadmium selenide (CdSe) nanorods.

This paper is organized as follows. The next section describes the hierarchical real-space grid DC-DFT algorithm for quantum mechanically based MD simulations and its parallelization. Simulation results are presented in Sec. III, and Sec. IV contains a summary.

II. DIVIDE-AND-CONQUER DENSITY-FUNCTIONAL-THEORY ALGORITHM ON HIERARCHICAL REAL-SPACE GRIDS

In the Kohn-Sham formulation[21] of the DFT,[20,22] the energy of a physical system is expressed as a functional of N atomic positions, $\vec{R}^N=\{R_I|I=1,\ldots,N\}$, and N_{band} wave functions (or Kohn-Sham orbitals), $\psi^{N_{\text{band}}}=\{\psi_n(\vec{r})|n=1,\ldots,N_{\text{band}}\}$:

$$E[\rho]=T_s[\rho]+\int d\vec{r}\rho(\vec{r})v_{\text{loc}}(\vec{r})+E_{\text{nl}}[\rho]+\frac{1}{2}\int d\vec{r}\int d\vec{r}'\frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}+E_{\text{xc}}[\rho]+\sum_{I<J}\frac{Z_IZ_J}{|\vec{R}_I-\vec{R}_J|}. \tag{1}$$

Here, $T_s[\rho]$ is the kinetic energy of a noninteracting electron gas in its ground state with the valence-electron number density,

$$\rho(\vec{r})=2\sum_{n=1}^{N_{\text{band}}}|\psi_n(\vec{r})|^2, \tag{2}$$

where only doubly occupied states are considered (we use the atomic unit). In Eq. (1),

$$v_{\text{loc}}(\vec{r})=\sum_I v_I^{\text{loc}}(\vec{r}-\vec{R}_I) \tag{3}$$

is a local ionic pseudopotential,[68]

$$E_{\text{nl}}[\rho]=\sum_n\sum_I\sum_{lm}\int d\vec{r}\psi_n^*(\vec{r})\xi_{lm}^I(\vec{r}-\vec{R}_I)\int d\vec{r}'\,\xi_{lm}^{I*}(\vec{r}'-\vec{R}_I)\psi_n(\vec{r}') \tag{4}$$

is a nonlocal pseudopotential energy,[67–69] where $\xi_{lm}^I(\vec{r}-\vec{R}_I)$ is a projection state localized at $\vec{R}_I$ with the angular momentum quantum numbers l and m, $E_{\text{xc}}[\rho]$ is the exchange-correlation energy, and the last term is the electrostatic energy between ions with valence $\{Z_I|I=1,\ldots,N\}$.

Given atomic positions $\vec{R}^N$, the ground-state energy is obtained by minimizing $E[\rho]$ with respect to $\psi^{N_{\text{band}}}$, subjected to orthonormality constraints,

$$\int d\vec{r}\psi_n^*(\vec{r})\psi_{n'}(\vec{r})=\delta_{nn'}\equiv\begin{cases}1 & (n=n')\\ 0 & (n\neq n').\end{cases} \tag{5}$$

The constrained minimization leads to the Kohn-Sham equations,

$$\hat{H}\psi_n(\vec{r})=\varepsilon_n\psi_n(\vec{r}), \tag{6}$$

where the Kohn-Sham Hamiltonian operator $\hat{H}$ is defined through

$$\hat{H}\psi_n(\vec{r})=\left[-\frac{1}{2}\nabla^2+v_{\text{loc}}(\vec{r})+v_{\text{H}}(\vec{r})+v_{\text{xc}}(\vec{r})\right]\psi_n(\vec{r})+\sum_I\sum_{lm}\xi_{lm}^I(\vec{r}-\vec{R}_I)\int d\vec{r}'\,\xi_{lm}^{I*}(\vec{r}'-\vec{R}_I)\psi_n(\vec{r}'). \tag{7}$$

In Eq. (7), ∇^2 is the Laplacian operator,

$$v_{\text{H}}(\vec{r})=\int d\vec{r}'\frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} \tag{8}$$

is the Hartree potential, and we use the generalized gradient approximation[70] to the exchange-correlation potential, $v_{\text{xc}}(\vec{r})=\delta E_{\text{xc}}/\delta\rho(\vec{r})$, which is a function of the local value of $\rho(\vec{r})$ and its gradient $\nabla\rho(\vec{r})$ at $\vec{r}$.

A. Divide-and-conquer density functional theory

We have developed an embedded divide-and-conquer (EDC) algorithmic framework[11] based on data locality principles to design linear-scaling algorithms for a wide range of

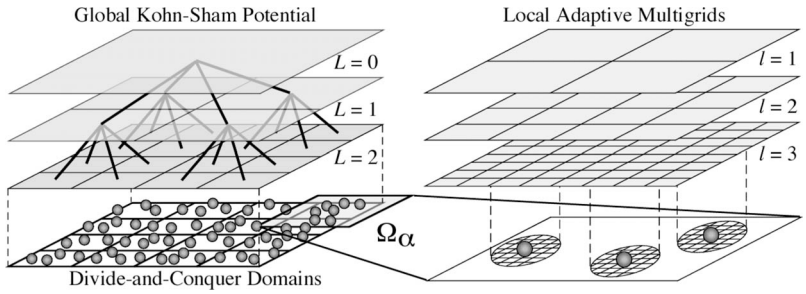

FIG. 1. Schematic of an embedded divide-and-conquer algorithm. (Left) The physical space is subdivided into spatially localized domains, with local atoms constituting subproblems (bottom), which are embedded in a global field (shaded) solved with a tree-based algorithm. In the case of the DFT, the global field is the Kohn-Sham potential. (Right) To solve the subproblem in domain Ω_α in the divide-and-conquer DFT algorithm, coarse multigrids (gray) are used to accelerate iterative solutions on the original real-space grid (corresponding to the grid refinement level, $l=3$). The bottom panel shows fine grids adaptively generated near the atoms (spheres) to accurately operate the ionic pseudopotentials on the electronic wave functions.

scientific applications with tight error control, including the DC-DFT algorithm in this paper. In EDC algorithms, spatially localized subproblems are solved in a global embedding field, which is efficiently computed with tree-based algorithms (Fig. 1). Examples of the embedding field are (1) the electrostatic field in MD simulation,[8] (2) the self-consistent Kohn-Sham potential in the DFT,[60] and (3) a coarser but less computer-intensive simulation method in hierarchical simulation, in which fine (e.g., DFT) simulations are embedded in coarse (e.g., MD) simulations on demand only when and where high accuracy is required.[29,71,72]

In the DC-DFT algorithm,[31–33,42,60] the three-dimensional space Ω is covered with overlapping domains (see Fig. 2),

$$\Omega = \bigcup_\alpha \Omega_\alpha. \tag{9}$$

Each domain Ω_α is further decomposed into its subvolumes,

$$\Omega_\alpha = \Omega_{0\alpha} \cup \Gamma_{1\alpha} \cup \Gamma_{2\alpha}, \tag{10}$$

where $\Omega_{0\alpha}$ is the nonoverlapping core,

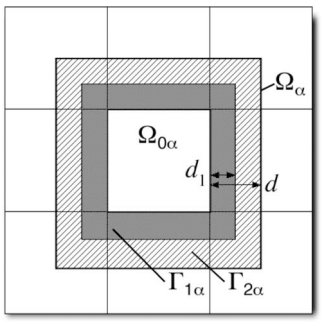

FIG. 2. Schematic of the divide-and-conquer algorithm in two dimensions. The physical space Ω is a union of overlapping domains, $\Omega = \cup_\alpha \Omega_\alpha$. Each domain is further decomposed into the nonoverlapping core $\Omega_{0\alpha}$, the primary buffer layer $\Gamma_{1\alpha}$ (shaded area), and the secondary buffer layer $\Gamma_{2\alpha}$ (hatched area). The width of the primary and total (=primary+secondary) buffer layers are d_1 and d, respectively.

$$\Omega = \bigcup_\alpha \Omega_{0\alpha}, \quad \Omega_{0\alpha} \cap \Omega_{0\beta} = 0 \quad (\alpha \neq \beta), \tag{11}$$

and $\Gamma_{1\alpha}$ and $\Gamma_{2\alpha}$ are the primary and secondary buffer layers, respectively.[33] The core $\Omega_{0\alpha}$ is enclosed with $\Gamma_{1\alpha}$, which in turn is surrounded by $\Gamma_{2\alpha}$ (Fig. 2).

For each domain α, we define a domain support function $p^\alpha(\vec{r})$ such that it is only nonzero within the core and primary buffer layer of the domain,

$$p^\alpha(\vec{r}) = 0 \quad \text{if } \vec{r} \notin \Omega_{0\alpha} \cup \Gamma_{1\alpha}, \tag{12}$$

with the sum rule,

$$\sum_\alpha p^\alpha(\vec{r}) = 1, \tag{13}$$

satisfied at every spatial position $\vec{r}$. Because of the sum rule, the valence-electron number density is exactly decomposed into

$$\rho(\vec{r}) = \sum_\alpha \rho^\alpha(\vec{r}), \tag{14}$$

where

$$\rho^\alpha(\vec{r}) = p^\alpha(\vec{r})\rho(\vec{r}) \tag{15}$$

is the partial contribution to the electron density from domain α.

The essential approximation in the DC-DFT algorithm is the replacement of the self-consistent Kohn-Sham Hamiltonian $\hat{H}$ by its subspace approximation $\hat{H}^\alpha$, which is formally identical to Eq. (7), but the Kohn-Sham equations,

$$\hat{H}^\alpha \psi_n^\alpha(\vec{r}) = \varepsilon_n^\alpha \psi_n^\alpha(\vec{r}), \tag{16}$$

are solved in each domain Ω_α to obtain locally orthonormal Kohn-Sham orbitals $\{\psi_n^\alpha(\vec{r})\}$:

$$\int_{\Omega_\alpha} d\vec{r} \psi_n^{\alpha *}(\vec{r}) \psi_{n'}^\alpha(\vec{r}) = \delta_{nn'} \equiv \begin{cases} 1 & (n = n') \\ 0 & (n \neq n'). \end{cases} \tag{17}$$

In Eq. (16), ε_n^α is the nth eigenvalue of $\hat{H}^\alpha$. Boundary conditions on $\{\psi_n^\alpha(\vec{r})\}$ are imposed at the domain boundary $\partial\Omega_\alpha$. We use either the rigid-wall boundary condition (in which

the wave function vanishes at $\partial\Omega_\alpha$) or the periodic boundary condition. The wave function values in the secondary buffer layer $\Gamma_{2\alpha}$ may be contaminated by the artificial boundary conditions at $\partial\Omega_\alpha$. This is the reason why the domain support function $p^\alpha(\vec{r})$ is made zero in $\Gamma_{2\alpha}$,[33] so that the contaminated wave function values do not contribute to the density $\rho(\vec{r})$. For the support function, we use a cubic interpolation function such that both the function value and its derivative are continuous at $p^\alpha(\vec{r})=0$ and 1.

To determine the number of occupied local Kohn-Sham orbitals in the DC-DFT algorithm, we first note that Eq. (2) is an expansion of the valence-electron density,

$$\rho(\vec{r}) = 2\langle\vec{r}|\Theta(\varepsilon_F - \hat{H})|\vec{r}\rangle, \tag{18}$$

in terms of the eigenstates of the Kohn-Sham Hamiltonian $\hat{H}$. In Eq. (18), $|\vec{r}\rangle$ is the coordinate eigenstate, the step function $\Theta(x)$ is 1 for $x>0$ and is 0 otherwise, and the Fermi energy ε_F is determined from the number of valence electrons $N_{\rm el}$ through the relation,

$$N_{\rm el} = \int d\vec{r}\rho(\vec{r}). \tag{19}$$

In the DC-DFT algorithm, the local Kohn-Sham orbitals $\{\psi_n^\alpha(\vec{r})\}$ are compactly supported on each domain Ω_α, and thus, the partial density $\rho^\alpha(\vec{r})$ is expanded in terms of them as

$$\rho^\alpha(\vec{r}) = p^\alpha(\vec{r})\sum_n f(\varepsilon_n^\alpha)|\psi_n^\alpha(\vec{r})|^2, \tag{20}$$

where $f(\varepsilon_n^\alpha)=2\Theta(\varepsilon_F-\varepsilon_n^\alpha)$. Accordingly, the normalization condition, Eq. (19), reads

$$N_{\rm el} = \sum_\alpha\sum_n f(\varepsilon_n^\alpha)\int d\vec{r}p^\alpha(\vec{r})|\psi_n^\alpha(\vec{r})|^2. \tag{21}$$

In DFT-based MD simulations, we numerically integrate Newton's equations of motion,

$$M_I\frac{d^2}{dt^2}\vec{R}_I = \vec{F}_I, \tag{22}$$

where the force $\vec{F}_I$ acting on the Ith ion is calculated from the Hellmann-Feynman theorem as

$$\vec{F}_I = \vec{F}_I^{\rm ion} + \vec{F}_I^{\rm loc} + \vec{F}_I^{\rm nl} = \sum_{J(\neq I)} Z_I Z_J \frac{\vec{R}_I - \vec{R}_J}{|\vec{R}_I - \vec{R}_J|^3} + \int d\vec{r}\rho(\vec{r})(\vec{r}-\vec{R}_I)\frac{dv_{\rm loc}}{d|\vec{r}-\vec{R}_I|} + \sum_\alpha\sum_n f(\varepsilon_n^\alpha)\sum_{lm}\left[\int d\vec{r}p^\alpha(\vec{r})\psi_n^{\alpha\,*}(\vec{r})\frac{\partial\xi_{lm}^I}{\partial(\vec{r}-\vec{R}_I)}\int d\vec{r}'\,\xi_{lm}^{I\,*}(\vec{r}' - \vec{R}_I)\psi_n^\alpha(\vec{r}') + \int d\vec{r}p^\alpha(\vec{r})\psi_n^{\alpha\,*}(\vec{r})\xi_{lm}^I(\vec{r}-\vec{R}_I)\int d\vec{r}'\frac{\partial\xi_{lm}^{I\,*}}{\partial(\vec{r}'-\vec{R}_I)}\psi_n^\alpha(\vec{r}')\right]. \tag{23}$$

In Eq. (23), the ionic contribution $\vec{F}_I^{\rm ion}$ can be computed efficiently with $O(N)$ operations using the fast multipole method (FMM).[73] Scalable FMM algorithms to compute $\vec{F}_I^{\rm ion}$ on parallel computers are described in our previous publications.[74,75] For systems with periodic boundary conditions, methods based on the Ewald summation are effective, including the $O(N\log N)$ particle mesh Ewald method.[76] (A parallel Ewald algorithm is described in Ref. 77.)

B. Hierarchical real-space grids

For efficient parallel implementation of DFT, we have developed a hierarchical real-space grid method based on higher-order finite differencing[61,62] and multigrid acceleration.[63,64] In the hierarchical-grid method, a real-space multigrid is adaptively refined[66] near each atom to accurately operate the ionic pseudopotentials on the electronic wave functions (see Fig. 1).

In the high-order finite-difference method for calculating the derivatives (i.e., kinetic-energy operator) in Eq. (7), $\psi_n^\alpha(\vec{r})$ and $\rho(\vec{r})$ are represented by numerical values on real-space grid points. The kinetic-energy operator is expanded using the finite-difference method as[61]

$$\left.\frac{\partial^2\psi_n(\vec{r})}{\partial x^2}\right|_{\vec{r}=(x_i,y_j,z_k)} = \sum_{\lambda=-L}^{L} C_\lambda\psi_n^\alpha(x_i+\lambda h, y_j, z_k) + O(h^{2L+2}), \tag{24}$$

where h is the grid spacing, and L is the order of the finite-difference method. Since the calculations are performed entirely in real space, this method is most suitable for spatially localized atomic configurations such as clusters, which is the case in the DC-DFT algorithm. Since the finite-difference expansion involves only short-ranged operations, an efficient implementation on parallel computers is possible.[37,38]

The constrained minimization of $E[\rho]$ with respect to $\{\psi_n^\alpha\}$ is performed iteratively based on the conjugate-gradient (CG) method.[78] The self-consistent minimization loop consists of the following steps:

(1) Calculate the electronic potentials, $v_{\rm loc}(\vec{r})$, $\{\xi_{lm}^I(\vec{r}-\vec{R}_I)\}$, $v_{\rm H}(\vec{r})$, and $v_{\rm xc}(\vec{r})$, using the density $\rho(\vec{r})$ from the previous step or the starting value (in the case of the first step). These potentials are common to all the domains; the Hartree potential is obtained by iterative solution to the Poisson equation, $\nabla^2 v_{\rm H}(\vec{r})=-4\pi\rho(\vec{r})$, using the multigrid method.[79]

(2) For all domains α

(a) Perform a unitary transformation of $\{\psi_n^\alpha\}$ to diagonalize the Hamiltonian matrix, $H_{nn'}^\alpha = \int_{\Omega_\alpha} d\vec{r}\psi_n^{\alpha*}(\vec{r})\hat{H}^\alpha\psi_{n'}^\alpha(\vec{r})$.

(b) Improve $\{\psi_n^\alpha(\vec{r})\}$ iteratively using the CG method.

(c) Orthonormalize $\{\psi_n^\alpha(\vec{r})\}$ according to Eq. (17) with the Gram-Schmidt method.[78]

(3) Determine the Fermi energy ε_F to reproduce the number of valence electrons $N_{\rm el}$ by solving Eq. (21) using the Newton-Raphson method.[79]

(4) Calculate the new density $\rho(\vec{r})$ according to Eqs. (14) and (20), using the updated $\{\psi_n^\alpha(\vec{r})\}$; refine $\rho(\vec{r})$ by mixing the new and old densities using the Pulay charge-mixing scheme.[80]

The steps 1–4 are repeated until the self-consistency between $\{\psi_n^\alpha\}$ and $\rho(\vec{r})$ is achieved within a prescribed error tolerance.

Step 2(b) in the self-consistent loop includes two inner loops: One is associated with the band index n and the other is the CG iteration for each band involving a loop index $i_{\rm CG}$. To reduce long wavelength components of the residual, we use the multigrid method.[5,38,63] On a coarser grid, we solve residual equations,

$$\left\{-\frac{1}{2}\nabla^2 + v(\vec{r})\right\}\varphi_n^\alpha(\vec{r}) = g_n^\alpha(\vec{r}), \tag{25}$$

where

$$g_n^\alpha(\vec{r}) = -(\hat{H}^\alpha - \varepsilon_n^\alpha)\psi_n^\alpha(\vec{r}) \tag{26}$$

is a residual vector. In this method, the electronic potential $v(\vec{r})$ as well as $g_n^\alpha(\vec{r})$ on a fine grid are restricted to a coarser grid using restriction operations. Here, the potential

$$v(\vec{r}) = v_{\rm loc}(\vec{r}) + v_{\rm H}(\vec{r}) + v_{\rm xc}(\vec{r}) \tag{27}$$

excludes the nonlocal pseudopotential, since it is short ranged and does not influence the solutions on coarser grids. The solution $\varphi_n^\alpha(\vec{r})$ of Eq. (25) on the coarse grid is prolongated to the fine grid and is added to $\psi_n^\alpha(\vec{r})$.

The computational complexity of the DC-DFT algorithm is $O(N)$ for N atoms, as analyzed below. Since the most time consuming part is step 2, we concentrate on the computational complexity of this step. The orthogonalization in step 2(c) requires $O(N_{\rm domain}n_{\rm band}^2 n_{\rm grid})$ operations, where $N_{\rm domain}$ is the number of domains, and $n_{\rm band}$ and $n_{\rm grid}$ are the number of Kohn-Sham orbitals and number of real-space grid points per domain, respectively. Since we choose $N_{\rm domain}=O(N)$, both $n_{\rm band}$ and $n_{\rm grid}$ are $O(n)=O(1)$, where $n=N/N_{\rm domain}$ is the average number of atoms per domain. Consequently, this operation scales as $O(Nn^3)=O(N)$. All calculations in step 2 except for the orthogonalization require $O(N_{\rm domain}n_{\rm band}n_{\rm grid})=O(Nn^2)=O(N)$ operations. The orthogonalization step, which scales as $O(Nn^3)$, is negligible[38] for relatively small domains ($n<100$) used in our calculations.

C. Parallelization

The hierarchical-grid DC-DFT algorithm has been implemented on parallel computers based on spatial decomposition,[8,38] in which the physical system (with lengths L_x, L_y, and L_z in the x, y, and z directions, respectively) is divided into $P=P_x\times P_y\times P_z$ rectangular subsystems of equal volume, and each subsystem is assigned to a processor in a parallel computer consisting of P processors (see Fig. 3). Specifically, atom I at position $\mathbf{R}_I=(R_{Ix},R_{Iy},R_{Iz})$ is assigned to processor $p\in[0,P-1]$, where

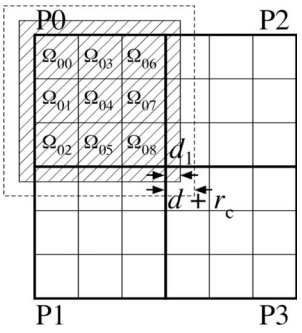

FIG. 3. Schematic of the parallel divide-and-conquer algorithm in two dimensions. The physical system is divided into subsystems, $P_0,\ldots,P_3$, of equal volume, and each subsystem is assigned to a processor in a parallel computer. Each subsystem, in turn, consists of multiple nonoverlapping core domains, $\Omega_{00},\ldots,\Omega_{08}$. To perform electronic-structure calculations on overlapping domains, $\Omega_{00},\ldots,\Omega_{08}$, on processor P_0, the contributions to the total electron density, $\rho(\vec{r})$, within the primary buffer-layer depth d_1 (see the hatched area), need to be cached from the nearest-neighbor processors. In addition, the ionic positions within width $d+r_c$ (enclosed by dashed lines) need to be cached from the nearest-neighbor processors to compute ionic pseudopotentials. Here, d and r_c are the total buffer-layer width and the range of nonlocal pseudopotentials, respectively.

$$p = p_x P_y P_z + p_y P_z + p_z,$$

$$p_\alpha = \lfloor R_{I\alpha}P_\alpha/L_\alpha \rfloor \quad (\alpha = x,y,z). \tag{28}$$

Each subsystem, in turn, consists of multiple domains.

In this parallelization scheme, each domain is local to a processor, and the Kohn-Sham orbitals $\{\psi_n^\alpha(\vec{r})\}$ need not be exchanged between processors. Consequently, there is no need for the massive communication required for orthonormalization in conventional parallel DFT algorithms[38]. Instead, there are two major sources of interprocessor communications in the parallel DC-DFT algorithm. First, for the computation of the total electron density $\rho(\vec{r})$ in a given domain Ω_α, partial contributions $\{\rho^{\alpha'}(\vec{r})\}$ within the primary buffer-layer depth d_1 need to be cached (i.e., copied by message exchanges) from the processors that contain the nearest-neighbor domains α'. Second, the positions of ions within depth $d+r_c$ need to be cached from the nearest-neighbor processors to compute ionic pseudopotentials. Here, d and r_c are the total buffer-layer width and the range of nonlocal pseudopotentials, respectively.

The computation time of the parallel DC-DFT algorithm scales as $O((N_{\rm domain}/P)n^3)=O((N/P)n^3)$. The communication time for the above two tasks, on the other hand, scales as $O((N/P)^{2/3})$. The communication overhead of this algorithm, $O((N/P)^{-1/3}n^{-3})$, is extremely low, because of the small

surface-to-volume ratio and the absence of communicating $O(n)$ Kohn-Sham orbitals. The parallel multigrid algorithm to compute the Hartree potential from $\rho(\vec{r})$ requires additional interprocessor communications that scale as $O(\log P)$.[81] For coarse-grained applications, in which $N_{\text{grid}}/P \gg P$ (N_{grid} is the total number of grid points), however, this logarithmic overhead is negligible.

The parallel DC-DFT program is written in FORTRAN 90 with message passing interface for message passing. Scalability tests of the parallel DC-DFT algorithm have been performed on a wide range of platforms, including the 131 072-processor IBM BlueGene/L at the Lawrence Livermore National Laboratory (LLNL), the 10 240-processor SGI Altix 3000 at the NASA Ames Research Center, and the 2048-processor AMD Opteron-based Linux cluster at the Collaboratory for Advanced Computing and Simulations of the University of Southern California (USC).

The BlueGene/L system at the LLNL comprises 65 536 computational node (CN) chips, each of which has two PowerPC 400 processors (131 072 processors in total) with 700 MHz clock speed. On single CN, the two processors share 512 Mbyte memory. Each processor has a 32 kbyte instruction/data cache, a 2 Mbyte L2 cache, and a 4 Mbyte L3 cache. The theoretical peak performance is 2.8 Gflops per processor. Two types of interconnection (three-dimensional torus and tree topologies) are designed for distinct purposes. The torus network is used mostly for point-to-point communications, whereas the tree network is optimized for collective communications. The interconnection bandwidths are 175 and 350 Mbyte/s per link, respectively, for the two networks.

The SGI Altix 3000 system named Columbia at the NASA Ames consists of 20 of SGI Altix model 3700 boxes, each consisting of 512 Intel 1.5 GHz Itanium2 processors. Each processor has 128 floating-point registers, a 32 kbyte L1 cache, a 256 kbyte L2 cache, and a 6 Mbyte L3 cache, and its theoretical peak performance is 6 Gflops. The SGI NUMAlink4 interconnection provides 1 Tbyte memory globally shared among cluster nodes within an Altix box. We have used up to four Altix boxes (up to 1920 processors) for our benchmark.

USC-HPCC operates 1824-node Linux cluster with 15.8 Tflops Linpack performance. We have used up to 512 nodes of dual-CPU dual-core AMD 2 GHz Opteron nodes (2048 processors in total). Each node has 4 Gbytes memory, and Myrinet interconnection provides 256 Mbyte/s bandwidth. Single Opteron core has 64 kbyte instruction/data caches and 1 Mbyte L2 cache.

Figure 4 shows the execution time per MD step as a function of the number of processors for the DC-DFT based MD algorithm to simulate alumina (Al_2O_3) systems on the Altix 3000. Here, the number of atoms N is scaled linearly with the number of processors P as $N=720P$. In the DC-DFT calculations, each domain of size $6.66\times5.76\times6.06$ Å^3 contains 40 electronic wave functions, where each wave function is represented on $28^3=21\,952$ grid points. The execution time includes three self-consistent (SC) iterations to determine the electronic wave functions and the Kohn-Sham potential, with three CG iterations per SC cycle to refine each wave function iteratively. The execution time increases only slightly as a function of P, which signifies an excellent parallel efficiency. We define the speed of a MD program as a product of the total number of atoms and time steps executed per second. The isogranular speedup is the ratio between the speed of P processors and that of one processor. The parallel efficiency is the speedup divided by P. The largest calculation on 1920 processors involves 1 382 400 atoms (121 385 779 200 electronic degrees of freedom), for which the isogranular parallel efficiency is 0.907. A better measure of the interbox scaling efficiency based on NUMAlink4 is the speedup from 480 processors in one box to 1920 processors in four boxes, divided by the number of boxes. On 1920 processors, the interbox scaling efficiency is 0.966. Also, the algorithm involves very small communication time (see Fig. 4).

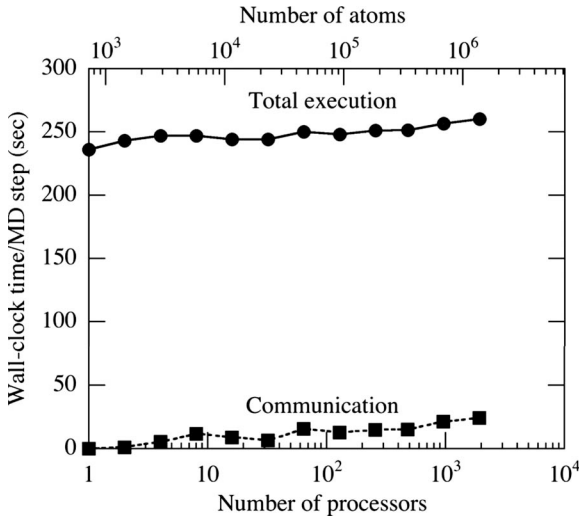

FIG. 4. Total execution (circles) and communication (squares) times per MD time step as a function of the number of processors P ($=1,\ldots,1920$) of Altix 3000 for the DC-DFT MD algorithm on $720P$ atom alumina systems.

Figure 5 shows the execution time of the DC-DFT algorithm per MD step as a function of the number of atoms for alumina systems on 131 072 BlueGene/L processors, 1920 Itanium2 processors of Altix 3000, and 2000 Opteron processors. On all three platforms, the execution time scales linearly with the number of atoms. The largest benchmark test in this study involves 11 796 480 atoms (1 035 825 315 840 electronic degrees of freedom) on 131 072 BlueGene/L processors. The floating-point performance of the DC-DFT algorithm on 1920 Itanium2 processors is 1.49 Tflops, compared to the theoretical peak performance of 11.5 Tflops.

III. RESULTS

We have applied the parallel DC-DFT algorithm to liquid rubidium (Rb), graphene, and cadmium selenide (CdSe) nanorods. In all these calculations, we use the norm-conserving pseudopotentials proposed by Troullier and Martins.[68] The generalized gradient approximation[70] is used for the exchange-correlation energy in the DFT. The grid spacing Δr in the real-space electronic-structure calculations is selected

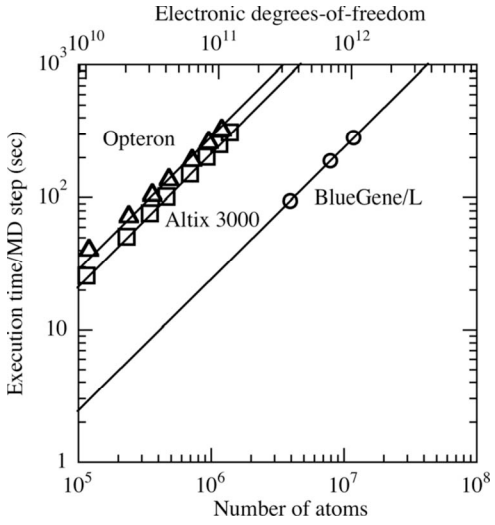

FIG. 5. Benchmark tests of the DC-DFT MD algorithm on 131 072 BlueGene/L processors (circles), 1920 Itanium2 processors of Altix 3000 at NASA (squares), and 2000 Opteron processors (triangles). The execution time per MD step is shown as a function of the number of atoms, and lines show $O(N)$ scaling.

to be small enough to obtain good convergence of the total energy of the system. Specifically, $\Delta r=1.00$, 0.36, and 0.54 a.u. for liquid Rb, graphene, and CdSe nanorod systems, respectively.

A. Liquid rubidium

The conventional DFT method has been applied extensively to liquid alkali metals[82–87] to study the correlation between conduction electrons and highly disordered state of ions. In an application of the DC-DFT algorithm to metallic systems, it is expected that the results are not sensitive to the selection of boundary if the buffer length is large enough, because there are no effects due to the covalent bond breaking and the existence of a dangling bond. Here, we use first-principles MD simulation with interatomic forces computed with the DC-DFT algorithm to study a disordered metallic system, i.e., liquid Rb. The purpose is to demonstrate the applicability of DC-DFT based MD simulation to metallic systems and to compare the calculated structure with those obtained experimentally as well as with the conventional DFT-based MD results. Regarding finite electronic temperature in metallic systems, both Fermi-Dirac and Gaussian-broadening schemes[88,89] to treat fractional occupation numbers of the local Kohn-Sham orbitals are implemented in our DC-DFT method. Since the Fermi energy ε_F is determined globally from the number of valence electrons N_{el} considering the local orbital energies in all domains, the fractional occupation numbers are calculated unambiguously, and the implementation is straightforward.

The DC-DFT based MD simulation involves 432 Rb atoms in a cubic cell of dimension 65.696 a.u. (Fig. 6). The size of the cell is determined from the experimental number density 0.0103 Å^{-3} at 350 K near the triple point. The squares shown by grid lines in Fig. 6 display the domains of length 16.424 a.u. The buffer size is 8.212 a.u. The MD

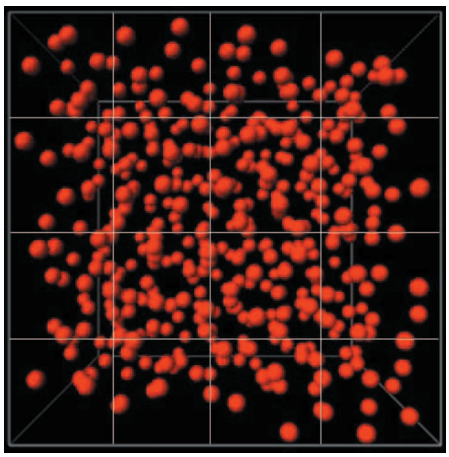

FIG. 6. (Color) Atomic configuration of liquid Rb. The grid lines show the domains in the DC-DFT calculations.

simulations are carried out at 350 K in the microcanonical ensemble. The equations of motion are integrated using the velocity Verlet algorithm with a time step of $\Delta t=4.8$ fs.

To test the applicability of the DC-DFT algorithm for first-principles MD simulations, Fig. 7 plots the total and potential energies as a function of time. The total energy is conserved within 2×10^{-4} a.u./ps per atom, which is 2 orders of magnitude smaller than the variation in the potential energy. (Before 2 ps, the velocities have been scaled to the prescribed values every 0.08 ps, and therefore, the total energy was not conserved.) Thus, the total energy is satisfactorily conserved in MD simulations.

Figure 8 compares the calculated pair-distribution functions $g(r)$ with a recent experimental result.[90] As shown in Fig. 8(a), the $g(r)$ obtained by the DC-DFT based MD simulation is in good agreement with the experiment for a wide range of distance r. In Fig. 8(b), $g(r)$ obtained by the conventional DFT-based MD simulation involving 54 Rb atoms[85] is compared with the experimental data. Because of the small system size, $g(r)$ for the 54-atom system is restricted to a limited range $\sim$10 Å. It is seen that the $g(r)$ near the cell boundary is affected by periodic boundary conditions, and, consequently, marked deviations from the experimental profile are recognized. On the other hand, $g(r)$ obtained by the DC-DFT algorithm has a smooth profile over the entire range of r in Fig. 8(a) without finite-size effects.

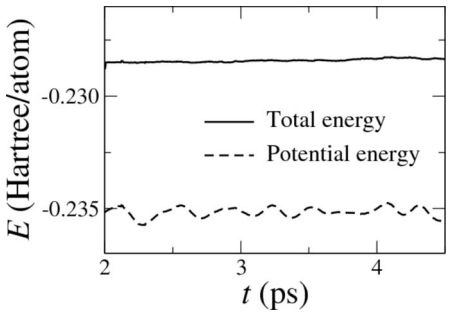

FIG. 7. Total (solid curve) and potential (dashed curve) energies in the atomic unit as a function of time in MD simulation of liquid Rb at 350 K (432 atoms in a cubic cell of side length 65.696 a.u.).

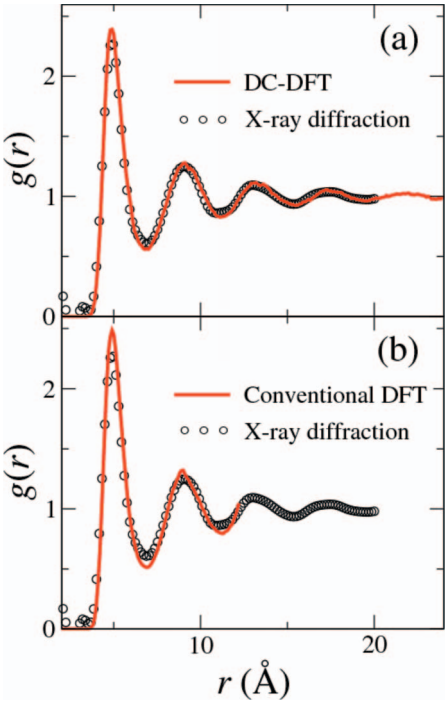

FIG. 8. (Color) Radial distribution functions $g(r)$ of liquid Rb at 350 K. The solid lines in (a) and (b) show 432-atom DC-DFT and 54-atom conventional DFT results, respectively. The open circles show the recent experimental results.

These results demonstrate the applicability of the DC-DFT algorithm for large-scale first-principles MD simulations of disordered metallic systems. Using this method, it is possible to study long-range density fluctuations near the critical point of liquid metals.

B. Graphene

For covalently bonded systems, the domains and boundaries should be carefully selected in the DC-DFT algorithm. To test the ability of the DC-DFT algorithm to covalent systems, we use a graphene sheet as shown in Fig. 9. The orthorhombic cell with dimensions of $(L_x, L_y, L_z) = (74.376\ \text{a.u.}, 64.398\ \text{a.u.}, 9.600\ \text{a.u.})$ contains 512 carbon atoms. The domain size is $9.297 \times 8.050 \times 9.600$ a.u..[3] The white grid lines in Fig. 9 show the domains in the xy plane.

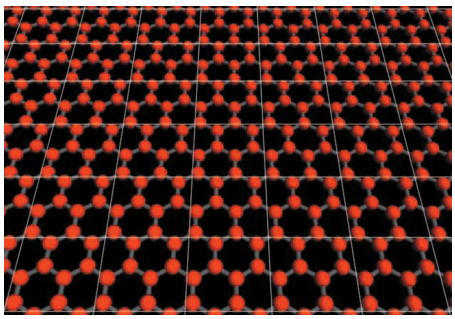

FIG. 9. (Color) Atomic configuration of graphene. The grid lines show the domains in the DC-DFT calculations.

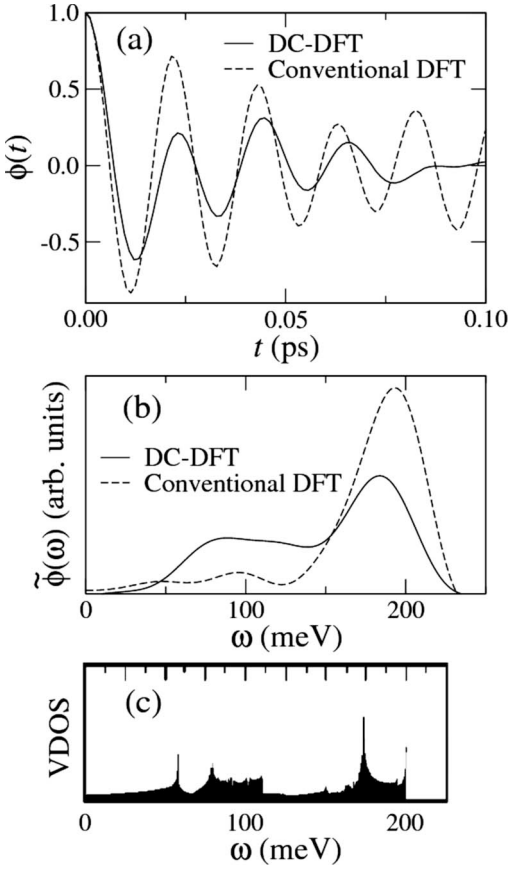

FIG. 10. (a) Velocity autocorrelation function $\phi(t)$ of graphene. (b) Fourier transformation $\tilde{\phi}(\omega)$ of $\phi(t)$. The solid and dashed lines show the 512-atom DC-DFT and 8-atom DFT results, respectively. (c) VDOS for graphene obtained by density-functional perturbation theory lattice dynamics in the generalized gradient approximation.

The buffer width in each direction is taken to be half the domain length. Microcanonical MD simulations are carried out at 300 K using the velocity Verlet algorithm with $\Delta t = 1.2$ fs. We find that the structure of graphene remains stable during the dynamics simulations. Since our selection of the domains and buffers is consistent with the periodicity of carbon atoms in the graphene, no covalent bond is "broken" at the boundaries of all domains when the electronic states are calculated locally. This helps stabilize the graphene structure.

To investigate the dynamic properties of graphene obtained by the DC-DFT based MD simulation, we calculate the velocity autocorrelation function $\phi(t)$ and its Fourier transformation $\tilde{\phi}(\omega)$, which are defined as

$$\phi(t) = \frac{\langle \vec{V}_I(t) \cdot \vec{V}_I(0) \rangle}{\langle \vec{V}_I(0) \cdot \vec{V}_I(0) \rangle} \quad \text{and} \quad \tilde{\phi}(\omega) = \int_0^\infty \phi(t) \cos \omega t dt,$$

respectively, where $\vec{V}_I(t)$ is the velocity of Ith atom at time t and $\langle \cdots \rangle$ denotes taking the average over both atoms and time origins. Figure 10 shows $\phi(t)$ and $\tilde{\phi}(\omega)$ (see the solid curves). For comparison, the dashed curves in Fig. 10 show conventional DFT-based MD results for a small system of eight carbon atoms. As shown by the solid curve in Fig.

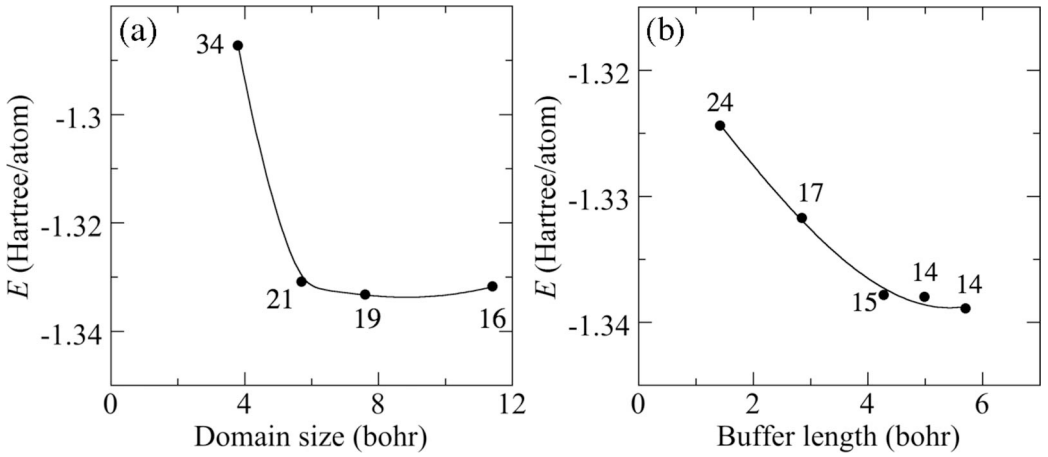

FIG. 11. (a) Potential energy per atom as a function of the domain size for an amorphous CdSe system (512 atoms in a cubic cell of side length 45.664 a.u.). The buffer size is fixed as 2.854 a.u. The atomic units are used for both energy and length. Numerals in the figure indicate the number of self-consistent iterations required for the convergence of the electron density within $\langle\{(\rho_i(\vec{r})-\rho_{i-1}(\vec{r}))/\rho_0\}^2\rangle \leq 10^{-4}$, where $\rho_i(\vec{r})$ is the electron density at ith iteration, ρ_0 is the average electron density, 0.0215 a.u., and the brackets denote the average over the grid in the entire system. (b) Potential energy as a function of the buffer width d for an amorphous CdSe system (512 atoms in a cubic cell of side length 45.664 a.u.). The domain size is fixed as 11.416 a.u. The atomic units are used for both energy and length. Numerals in the figure indicate the number of self-consistent iterations required for the convergence of the electron density.

10(b), $\tilde{\phi}(\omega)$ for the 512-atom system correctly describes lower-frequency components with $\omega < 140$ meV, which corresponds to the long-range acoustic modes of graphene, while $\tilde{\phi}(\omega)$ for the 8-atom system consists of only the higher-frequency peak, which corresponds to the short-range optical modes. The frequency range and peak positions for the large system are consistent with the vibrational density of states[91] (VDOS) obtained by density functional perturbation theory lattice dynamics displayed in Fig. 10(c) and high energy electron-energy loss spectroscopy measurements.[92]

C. Cadmium selenide

It is not obvious whether the DC-DFT algorithm is applicable to systems with broken covalent bonds such as amorphous semiconductors and clusters. To test the applicability, we calculate the band gaps (BGs) of semiconductor nanocrystals with the DC-DFT algorithm and compare them with those obtained by a conventional DFT method. The BG is calculated as BG=IP−EA, where IP and EA are the ionization potential and the electron affinity, respectively.[93] Since our formalism is completely in the real space, it is easy to obtain the IP and EA for finite systems by extracting and adding one electron, respectively. However, it is necessary to specify the boundary where the amplitude of electronic wave functions becomes zero, especially for systems with an extra electron. To obtain the electronic wave functions, we use a spheroidal boundary, and specify it by the distance R between the outermost atom and the boundary. To determine the BG, it is crucial to know the R dependence of the BG, because it is expected that both the IP and EA depend strongly on R.

Before simulating CdSe nanorods, we tested the convergence properties of the DC-DFT algorithm. For these numerical tests and subsequent simulations, we used periodic boundary conditions to obtain local Kohn-Sham orbitals, because a faster convergence was achieved compared with the rigid-wall boundary conditions. Since we discard the wave functions within the secondary buffer layers, numerical results are likely to be insensitive to the choice of the support function.

The DC-DFT algorithm has a well-defined set of localization parameters, with which the computational cost and the accuracy are controlled, i.e., the size of a domain and the length of buffer layers. We first test the convergence of the algorithm with respect to the domain size. Figure 11(a) shows the potential energy as a function of the domain size for an amorphous CdSe system containing 512 atoms in a cubic cell of length 45.664 a.u. The amorphous configuration is prepared by a melt-quench procedure[94] in MD simulation based on an empirical interatomic potential.[95] Here, and in the following, the grid spacing is chosen as 0.476 a.u. The total buffer size d is fixed at 2.854 a.u. In all the numerical tests, the primary buffer size is chosen as $d_1=d/2$. The energy converges within 0.003 a.u. per atom above the domain size of 6 a.u. The number of self-consistent iterations required for convergence decreases with increased domain size, as shown by the numerals in Fig. 11(a).

Next, we test the convergence of the potential energy as a function of the buffer size d for an amorphous CdSe system containing 512 atoms in a cubic cell of length 45.664 a.u. [see Fig. 11(b)]. The domain size of 11.416 a.u. is fixed. The potential energy converges within 0.001 a.u. per atom above $d=4$ a.u. The number of self-consistent iterations required for convergence decreases with increased buffer length size, as shown by the numerals in Fig. 11(b).

We use a CdSe nanorod[95,96] consisting of 216 Cd and 216 Se atoms. The atomic configuration of the nanorod is shown in Fig. 12. The diameter and length of the rod are 35.8 and 54.0 a.u., respectively. In the DC-DFT calculations, the domain and buffer lengths are chosen as ~8.0 and 6.0 a.u.,

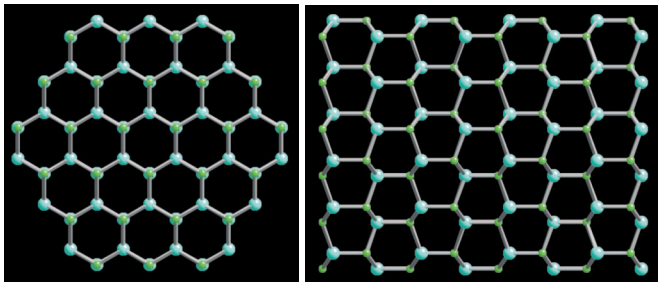

FIG. 12. (Color) Top (left panel) and side (right panel) views of a $Cd_{216}Se_{216}$ nanorod. The green and blue spheres show the positions of Cd and Se atoms, respectively.

respectively, which are slightly changed depending on R because of geometric reasons.

Figure 13 shows the IP, EA, and BG obtained by both the DC-DFT and conventional DFT methods. The agreement between the two results is excellent (within 0.01 hartree at $R=14.5$ a.u.) where the latter is available, which indicates that the intrinsic properties of the electronic states of CdSe nanorods are described correctly by the DC-DFT method. The DC-DFT algorithm allows the calculation to be extended to larger R, where the BG value converges to 1.1 eV above $R=20$ a.u. Experimentally, the BG of rodlike CdSe quantum dots has been measured to be ~ 2 eV depending on the size of the quantum dots.[97] The discrepancy between the theoretical and experimental values may be due to the treatment of the surface. Passivating the surface by hydrogenlike atoms may increase the theoretical BG as pointed out by Chelikowsky and co-workers.[98,99]

IV. SUMMARY

In conclusion, we have developed a linear-scaling divide-and-conquer density-functional-theory algorithm for large first-principles molecular-dynamics simulations. In the hierarchical-grid DC-DFT algorithm, electronic wave functions are represented on a hierarchical real-space grid that combines a coarse multigrid and adaptive fine grids around atoms. With spatial decomposition, the DC-DFT algorithm has been implemented on massively parallel computers with high parallel efficiency, e.g., 11.8×10^6-atom (1.04×10^{12} electronic degrees of freedom) calculation on 131 072 IBM BlueGene/L processors. The DC-DFT algorithm has well-defined parameters to control the data locality, with which

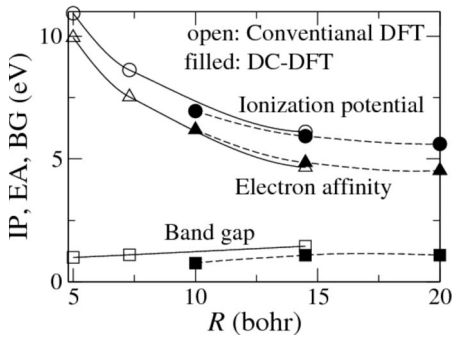

FIG. 13. Boundary length R dependences of IP, EA, and BG of a $Cd_{216}Se_{216}$ nanorod. The circles, triangles, and squares show IP, EA, and BG, respectively. The filled and open symbols show the results obtained by the DC and conventional DFT methods, respectively.

the solutions converge rapidly. Also, the total energy is well conserved during MD simulation. We have performed first-principles MD simulations based on the DC-DFT algorithm, in which the large system sizes have brought in excellent agreement with x-ray scattering measurements of the pair-distribution function of liquid Rb and have allowed correct description of low-frequency vibrational modes of graphene. The band gap of CdSe nanorods calculated by the DC-DFT algorithm agrees well with conventional DFT results. With the DC-DFT algorithm, we have calculated the band gap for larger system sizes until it saturates to an asymptotic value.

ACKNOWLEDGMENTS

This work was partially supported by ARO MURI's, DTRA, DOE-BES, DOE-SciDAC, NSF-ITR, and NSF-PetaApps. F.S. acknowledges support by Grant-in-Aid for Scientific Research on Priority Area, "Nanoionics (439)" from the MEXT, Japan. Simulations were performed at the University of Southern California using the 5384-processor Linux cluster at the Research Computing Facility and the 2048-processor Linux cluster at the Collaboratory for Advanced Computing and Simulations. We thank R. Biswas, K. Nomura, D. Srivastava, and L. H. Yang for their help with scalability tests on the IBM BlueGene/L computer at the Lawrence Livermore National Laboratory and the SGI Altix 3000 computer at the NASA Ames Research Center.

[1] J. J. Gilman, Science **274**, 65 (1996).
[2] D. W. Brenner, D. H. Robertson, M. L. Elert, and C. T. White, Phys. Rev. Lett. **70**, 2174 (1993).
[3] A. Strachan, A. van Duin, D. Chakraborty, S. Dasgupta, and W. A. Goddard, Phys. Rev. Lett. **91**, 098301 (2003).
[4] Ken-ichi Nomura, Rajiv K. Kalia, Aiichiro Nakano, Priya Vashishta, Adri C. T. van Duin, and William A. Goddard III, Phys. Rev. Lett., **99**, 148303 (2007).
[5] S. Ogata, F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, J. Appl. Phys. **95**, 5316 (2004).
[6] F. F. Abraham, R. Walkup, H. J. Gao, M. Duchaineau, T. D. De la Rubia, and M. Seager, Proc. Natl. Acad. Sci. U.S.A. **99**, 5777 (2002).
[7] T. C. Germann and P. S. Lomdahl, Comput. Sci. Eng. **1**, 10 (1999).
[8] A. Nakano, R. K. Kalia, P. Vashishta, T. J. Campbell, S. Ogata, F.

Shimojo, and S. Saini, Sci. Prog. **10**, 263 (2002).
[9] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[10] J.-L. Fattebert and F. Gygi, Comput. Phys. Commun. **162**, 24 (2004).
[11] A. Nakano, R. K. Kalia, K. Nomura, A. Sharma, P. Vashishta, F. Shimojo, A. C. T. v. Duin, I. W. A. Goddard, R. Biswas, and D. Srivastava, Comput. Mater. Sci. **38**, 642 (2007).
[12] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, J. Phys. Chem. A **105**, 9396 (2001).
[13] J. Tersoff, Phys. Rev. B **37**, 6991 (1988).
[14] D. W. Brenner, Phys. Status Solidi B **217**, 23 (2000).
[15] A. K. Rappe and W. A. Goddard, J. Phys. Chem. **95**, 3358 (1991).
[16] F. H. Streitz and J. W. Mintmire, Phys. Rev. B **50**, 11996 (1994).
[17] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).
[18] A. Nakano, Comput. Phys. Commun. **104**, 59 (1997).
[19] T. Campbell, R. K. Kalia, A. Nakano, P. Vashishta, S. Ogata, and S. Rodgers, Phys. Rev. Lett. **82**, 4866 (1999).
[20] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[21] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[22] W. Kohn and P. Vashishta, *Inhomogeneous Electron Gas*, edited by N. H. March and S. Lundqvist (Plenum, New York, 1983), Vol. 79.
[23] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
[24] W. E. Pickett, Comput. Phys. Rep. **9**, 115 (1989).
[25] G. Kresse and J. Furthmuller, Phys. Rev. B **54**, 11169 (1996).
[26] M. E. Tuckerman, D. A. Yarne, S. O. Samuelson, A. L. Hughes, and G. J. Martyna, Comput. Phys. Commun. **128**, 333 (2000).
[27] B. B. Karki, R. M. Wentzcovitch, S. de Gironcoli, and S. Baroni, Science **286**, 1705 (1999).
[28] H. Kikuchi, R. K. Kalia, A. Nakano, P. Vashishta, H. Iyetomi, S. Ogata, T. Kouno, F. Shimojo, K. Tsuruta, and S. Saini, *Proceedings of Supercomputing 2002* (IEEE, New York, 2002).
[29] S. Ogata, E. Lidorikis, F. Shimojo, A. Nakano, P. Vashishta, and R. K. Kalia, Comput. Phys. Commun. **138**, 143 (2001).
[30] S. Ogata, F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, Comput. Phys. Commun. **149**, 30 (2002).
[31] W. T. Yang, Phys. Rev. Lett. **66**, 1438 (1991).
[32] W. T. Yang and T. S. Lee, J. Chem. Phys. **103**, 5674 (1995).
[33] S. L. Dixon and K. M. Merz, J. Chem. Phys. **107**, 879 (1997).
[34] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[35] G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992).
[36] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[37] F. Shimojo, T. J. Campbell, R. K. Kalia, A. Nakano, P. Vashishta, S. Ogata, and K. Tsuruta, FGCS, Future Gener. Comput. Syst. **17**, 279 (2000).
[38] F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, Comput. Phys. Commun. **140**, 303 (2001).
[39] C. K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, Phys. Status Solidi B **243**, 973 (2006).
[40] D. R. Bowler, R. Choudhury, M. J. Gillan, and T. Miyazaki, Phys. Status Solidi B **243**, 989 (2006).
[41] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon, and D. Sanchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).
[42] T. Ozaki, Phys. Rev. B **74**, 245101 (2006).
[43] X. P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).
[44] M. S. Daw, Phys. Rev. B **47**, 10895 (1993).
[45] J. N. Kim, F. Mauri, and G. Galli, Phys. Rev. B **52**, 1640 (1995).
[46] P. Ordejón, D. A. Drabold, R. M. Martin, and M. P. Grumbach, Phys. Rev. B **51**, 1456 (1995).
[47] E. B. Stechel, A. R. Williams, and P. J. Feibelman, Phys. Rev. B **49**, 10088 (1994).
[48] A. J. Williamson, R. Q. Hood, and J. C. Grossman, Phys. Rev. Lett. **87**, 246406 (2001).
[49] G. E. Scuseria and P. Y. Ayala, J. Chem. Phys. **111**, 8330 (1999).
[50] J. E. Subotnik, A. Sodt, and M. Head-Gordon, J. Chem. Phys. **125**, 074116 (2006).
[51] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[52] M. Challacombe and E. Schwegler, J. Chem. Phys. **106**, 5526 (1997).
[53] E. Tsuchida, J. Phys. Soc. Jpn. **76**, 034708 (2007).
[54] C. Y. Yam, S. Yokojima, and G. H. Chen, Phys. Rev. B **68**, 153105 (2003).
[55] N. Bernstein, Europhys. Lett. **55**, 52 (2001).
[56] M. Geshi, T. Hoshi, and T. Fujiwara, J. Phys. Soc. Jpn. **72**, 2880 (2003).
[57] H. Y. Liu, M. Elstner, E. Kaxiras, T. Frauenheim, J. Hermans, and W. T. Yang, Proteins: Struct., Funct., Genet. **44**, 484 (2001).
[58] Y. Wang, G. M. Stocks, W. A. Shelton, D. M. C. Nicholson, Z. Szotek, and W. M. Temmerman, Phys. Rev. Lett. **75**, 2867 (1995).
[59] E. Prodan and W. Kohn, Proc. Natl. Acad. Sci. U.S.A. **102**, 11635 (2005).
[60] F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, Comput. Phys. Commun. **167**, 151 (2005).
[61] J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, Phys. Rev. B **50**, 11355 (1994).
[62] J. R. Chelikowsky, Y. Saad, S. Ögüt, I. Vasiliev, and A. Stathopoulos, Phys. Status Solidi B **217**, 173 (2000).
[63] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **54**, 14362 (1996).
[64] J.-L. Fattebert and J. Bernholc, Phys. Rev. B **62**, 1713 (2000).
[65] T. L. Beck, Rev. Mod. Phys. **72**, 1041 (2000).
[66] T. Ono and K. Hirose, Phys. Rev. Lett. **82**, 5016 (1999).
[67] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[68] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 8861 (1991).
[69] R. D. King-Smith, M. C. Payne, and J. S. Lin, Phys. Rev. B **44**, 13063 (1991).
[70] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[71] A. Nakano, M. E. Bachlechner, R. K. Kalia, E. Lidorikis, P. Vashishta, G. Z. Voyiadjis, T. J. Campbell, S. Ogata, and F. Shimojo, Comput. Sci. Eng. **3**, 56 (2001).
[72] H. Takemiya, Y. Tanaka, S. Sekiguchi, S. Ogata, R. K. Kalia, A. Nakano, and P. Vashishta, *Proceedings of Supercomputing 2006* (IEEE/ACM, New York, 2006).
[73] L. Greengard and V. Rokhlin, J. Comput. Phys. **73**, 325 (1987).
[74] A. Nakano, R. K. Kalia, and P. Vashishta, Comput. Phys. Commun. **83**, 197 (1994).
[75] S. Ogata, T. J. Campbell, R. K. Kalia, A. Nakano, P. Vashishta, and S. Vemparala, Comput. Phys. Commun. **153**, 445 (2003).
[76] T. Darden, D. York, and L. Pederson, J. Chem. Phys. **98**, 10089 (1993).
[77] R. K. Kalia, A. Nakano, D. L. Greenwell, P. Vashishta, and S. W. d. Leeuw, Supercomputer **54**, 11 (1993).
[78] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D.

Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[79] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 2nd ed. (Cambridge University Press, Cambridge, England, 1992).
[80] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).
[81] A. Nakano, P. Vashishta, and R. K. Kalia, Comput. Phys. Commun. **83**, 181 (1994).
[82] G. X. Qian, M. Weinert, G. W. Fernando, and J. W. Davenport, Phys. Rev. Lett. **64**, 1146 (1990).
[83] D. M. Bylander and L. Kleinman, Phys. Rev. B **45**, 9663 (1992).
[84] M. Foley, E. Smargiassi, and P. A. Madden, J. Phys.: Condens. Matter **6**, 5231 (1994).
[85] F. Shimojo, Y. Zempo, K. Hoshino, and M. Watabe, Phys. Rev. B **52**, 9320 (1995).
[86] R. M. Wentzcovitch and J. L. Martins, Solid State Commun. **78**, 831 (1991).
[87] R. M. Wentzcovitch, J. L. Martins, and P. B. Allen, Phys. Rev. B **45**, 11372 (1992).
[88] M. Weinert and J. W. Davenport, Phys. Rev. B **45**, 13709 (1992).
[89] C. Elsasser, M. Fahnle, C. T. Chan, and K. M. Ho, Phys. Rev. B **49**, 13975 (1994).
[90] K. Matsuda, K. Tamura, and M. Inui, Phys. Rev. Lett. **98**, 096401 (2007).
[91] N. Mounet and N. Marzari, Phys. Rev. B **71**, 205214 (2005).
[92] H. Yanagisawa, T. Tanaka, Y. Ishida, M. Matsue, E. Rokuta, S. Otani, and C. Oshima, Surf. Interface Anal. **37**, 133 (2005).
[93] A. Gorling and M. Levy, Phys. Rev. A **52**, 4493 (1995).
[94] J. P. Rino, I. Ebbsjo, P. S. Branicio, R. K. Kalia, A. Nakano, F. Shimojo, and P. Vashishta, Phys. Rev. B **70**, 045207 (2004).
[95] N. J. Lee, R. K. Kalia, A. Nakano, and P. Vashishta, Appl. Phys. Lett. **89**, 093101 (2006).
[96] J. T. Hu, L. S. Li, W. D. Yang, L. Manna, L. W. Wang, and A. P. Alivisatos, Science **292**, 2060 (2001).
[97] L. S. Li, J. T. Hu, W. D. Yang, and A. P. Alivisatos, Nano Lett. **1**, 349 (2001).
[98] X. Y. Huang, E. Lindgren, and J. R. Chelikowsky, Phys. Rev. B **71**, 165328 (2005).
[99] G. M. Dalpian and J. R. Chelikowsky, Phys. Rev. Lett. **96**, 226802 (2006).