Cell multipole method for molecular simulations in bulk and confined systems

Jie Zheng
Department of Chemical Engineering, University of Washington, Seattle, Washington 98195

Ramkumar Balasundaram and Stevin H. Gehrke
Department of Chemical Engineering, Kansas State University, Manhattan, Kansas 66506

Grant S. Heffelfinger
Computational Biology and Materials Technology Department, Sandia National Laboratories, Albuquerque, New Mexico 87185

William A. Goddard III
Materials and Process Simulation Center, Beckman Institute, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Shaoyi Jiang[a)]
Department of Chemical Engineering, University of Washington, Seattle, Washington 98195

(Received 25 February 2002; accepted 30 December 2002)

One of the bottlenecks in molecular simulations is to treat large systems involving electrostatic interactions. Computational time in conventional molecular simulation methods scales with $O(N^2)$, where N is the number of atoms. With the emergence of new simulations methodologies, such as the cell multipole method (CMM), and massively parallel supercomputers, simulations of 10-million atoms or more have been performed. In this work, the optimal hierarchical cell level and the algorithm for Taylor expansion were recommended for fast and efficient molecular dynamics simulations of three-dimensional (3D) systems. CMM was then extended to treat quasi-two-dimensional (2D) systems, which is very important for condensed matter physics problems. In addition, CMM was applied to grand canonical ensemble Monte Carlo simulations for both 3D and 2D systems. Under the optimal conditions, our results show that computational time is approximately linear with N for large systems, average error in total potential energy is about 0.05% for 3D and 0.32% for 2D systems, and the RMS force error is 0.27% for 3D and 0.43% for 2D systems when compared with the Ewald summation. © *2003 American Institute of Physics.*
[DOI: 10.1063/1.1553979]

I. INTRODUCTION

The fast and accurate treatment of long-range Coulombic interactions for large systems is one of the most challenging tasks in computer simulations of charged particles. For a three-dimensional (3D) periodic system, the Ewald summation method (EW3D) has been widely used to handle long-range electrostatic interactions between charged particles. However, the Ewald summation method is computationally very expensive since its complexity is $O(N^{1.5})$ in an N-particle system. A common approach is to truncate the interactions at a certain cutoff distance. This reduces operation count to $O(N)$, but significantly sacrifices accuracy, particularly for long-range Coulombic interactions.[1] Recently, much effort has been devoted to improving the efficiency of the Ewald summation method and developing alternative methods for large systems, such as the particle–particle and particle–mesh method (PPPM)[4,5] and the cell multipole method (CMM).[1–3] Both PPPM and CMM are more efficient than the Ewald summation. CMM is well suited to massively parallel supercomputers due to its hierarchical tree structure. In particular, its multipole-based and hierarchical cell structures are well suited to calculating long-range interactions in large systems. The cost is reduced form $O(N^2)$ to $O(N \log N)$. With growing interest in surface and interfacial systems, it is desirable to apply CMM to quasi-two-dimensional (2D) systems, where periodicity exists in only two dimensions.

In this work, the optimal hierarchical cell (e.g., level 2 or 3) and the algorithm for Taylor expansion (e.g., complex or simple downward) were recommended for fast and efficient molecular dynamics (MD) simulations of 3D systems. CMM was then extended to MD simulations for quasi-2D systems and grand canonical ensemble Monte Carlo (GCMC) simulations for both 3D and quasi-2D systems. These methods were tested on pure, binary, and ternary systems. This paper is organized as follows. In the next section we describe the basic ideas of CMM and the reduced cell multipole method (RCMM) for periodic systems in MD simulations, and the extension of CMM to confined systems and to GCMC simulations. In the third section, various simulation results of pure

a)Author to whom correspondence should be addressed. Electronic mail: sjiang@u.washington.edu

liquids or mixtures (i.e., water, methane, and dendrimers) in bulk and confined systems by MD and GCMC simulations are given. Conclusions are presented in the final section.

II. METHODOLOGY

A. Multipole approximation

The key idea of CMM is that by dividing a system into cubic cells and using multipoles (i.e., charges, dipoles, and quadrupoles) to represent these cells, a large number of atoms that lead to $O(N^2)$ computations for calculating long-range interactions in the far-field are replaced by these multiple moments.[6,7] This replacement reduces computational time to $O(N \log N)$. In order to maintain accuracy, nearby nonbonded interactions are computed explicitly while distant interactions are evaluated by multipoles and Taylor expansions.

CMM is used for any inverse power-law interaction potential[2] as $V=\Sigma_{i>j} q_i q_j / r_{ij}^p$, where $p=1$ for Coulombic, $p=6$ for Lennard-Jones (LJ) dispersive, and $p=12$ for LJ repulsive interactions. CMM can be divided into four parts.[1]

(i) *Cell decomposition.* The simulation system is decomposed into a hierarchy of cells like a tree structure. The root of the tree is the original system that is defined as level 0 while the leaf of the tree is at the deepest level. The effect of each cell will be represented by multipole expansions.

(ii) *Multipole expansion.* The multipole moments for all the cells at the deepest level are first determined by using[25]

$$Z=\sum_i q_i, \tag{1a}$$

$$\mu_\alpha=\sum_i p q_i r_{i\alpha}, \tag{1b}$$

$$Q_{\alpha\beta}=\frac{1}{2}\sum q_i[p(p+2)r_{i\alpha}r_{i\beta}-p\,\delta_{\alpha\beta}r_i^2], \tag{1c}$$

where Z, μ_α, $Q_{\alpha\beta}$ are the charge, dipole, and quadrupole. r_i is the distance between atom i and the center of the cell containing atom i. α and β are the x,y,z components in the Cartesian system. Then, the multipoles at the higher level are calculated by combining the corresponding multipole moments of eight child cells at the lower level. This process starts from the deepest level, and then moves upward to the root level (level 0). Thus, for a given atom, the potential energy from the far-field is evaluated by multipole moment expansions:

$$E_{\text{for}}=\frac{Z}{R^p}+\frac{\mu_\alpha R_\alpha}{R^{p+2}}+\frac{Q_{\alpha\beta}R_\alpha R_\beta}{R^{p+4}}, \tag{2}$$

where R is the distance between atom i and the center of the cell in the far-field.

(iii) *Far-field multipole Taylor expansion*. Since the multipole moments in the far-field are all the same for each atom in a given leaf cell, it is reasonable to expand the multipole moments in terms of a series of Taylor coefficients from all distant cells in the far-field to the center of the given cell instead of to each atom. This saves computational time. Thus, Eq. (2) is changed to

$$E_{\text{far}}=\xi^{(0)}+\xi_\alpha^{(1)}r_\alpha+\xi_{\alpha\beta}^{(2)}r_\alpha r_\beta, \tag{3}$$

where $\xi^{(i)}$ is the summation of ith order of Taylor expansion coefficients from multipole moments. r_α is the distance between atom i and the center of the cell containing atom i. As shown in Fig. 1, for one of eighty gray cell (in three dimensions), the simple downward algorithm involved Taylor expansion coefficients $\xi^{(i)}$, which are calculated by directly summing all the contributions from far-field cells marked as 1 and 2. However, as we can see in Fig. 1, eight gray cells, which come from the same parent cell, have exactly the same distant cells 2, but different cells 1. Thus, an alternative efficient method for calculating the coefficients of the Taylor expansion around the center of each gray cell is the so-called complex downward algorithm. In this algorithm, Taylor expansion is first established around the center of the parent cell from distant cells 2. Taylor expansion around the center of each child cell consists of two contributions—by Taylor expansion around the center of each child cell directly from distant cells 1 and by shifting Taylor expansion coefficients around the center of the parent center obtained from distant cells 2 in the center of each child cell using the following equations:

$$\begin{aligned}
\xi_{\text{gray}}^{(0)}&=\xi_1^{(0)}+\xi_2^{(0)}+\xi_{2,\alpha}^{(1)}r'+\xi_{2,\alpha\beta}^{(2)}r'_\alpha r'_\beta,\\
\xi_{\text{gray},\alpha}^{(1)}&=\xi_{1,\alpha}^{(1)}+\xi_{2,\alpha}^{(1)}+2\,\xi_{2,\alpha\beta}^{(2)}r'_\beta,\\
\xi_{\text{gray},\alpha\beta}^{(2)}&=\xi_{1,\alpha\beta}^{(2)}+\xi_{2,\alpha\beta}^{(2)},
\end{aligned} \tag{4}$$

where ξ_1 and ξ_2 are the Taylor coefficients from distant cells 1 and 2, respectively. r' is the distance between the centers of the parent and child cells.

(iv) *Near-field and far-field computations* For a given leaf cell, it has 27 neighbor cells including itself and the interactions in this near-field are computed explicitly in terms of $E_{\text{near}}=1/2\Sigma_i\Sigma_{j\neq i}(q_i q_j / r_{ij}^p)$. The remaining far-field interactions E_{far} are evaluated by multipole moment expansions.

RCMM is a relatively simple way of extending CMM to periodic systems. The most difficult problem with infinite crystals is to compute Coulombic interactions, $V_i=\Sigma_i' q_j / r_{ij}$, which are conditionally convergent.[3] The key idea of RCMM is that the original unit cell (containing 100 or 10 million atoms) is replaced by a reduced cell containing a small number of randomly positioned fictitious charges (e.g., 35 in this case), which can reproduce up to 5th multipole moments of the unit cell. We solve 35 equations [Eq. (1)] from the known multipoles (up to 5th multipole moments) of the unit cell for 35 fictitious charges. The difference between the reduced cell and the original cell is at the kth ($k=5$) multipole moments, which fall off very fast as $1/r^{K+1}$. In RCMM, Coulombic interactions are divided into two terms:[8] (a) the potential generated by charges in the original unit cell and 26 nearest images neighbor cells, and (b) the potential due to the fictitious charges in the remaining $\infty-27$ cells. The first term (a) is calculated with CMM, and the second term (b) involves the Ewald summation over fictitious atoms.[8] For short-range van der Waals (VDW) interactions, RCMM is not needed.

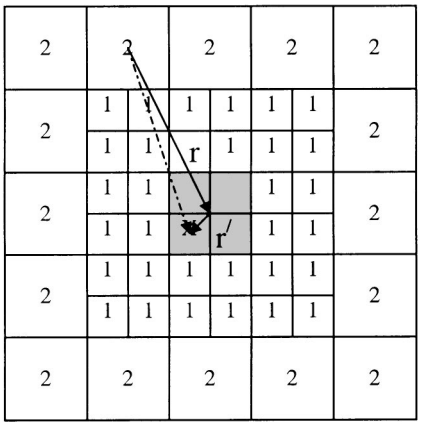

FIG. 1. For the simple and complex downward, Taylor expansion is established around the center of each child cell (dashed line) and of the parent cell (solid line) from distant cells 2, respectively. For the complex downward, Taylor expansion around the center of each child cell is further established based on the coefficients of the Taylor expansion around the center of the parent cell.

B. Extension to 2D systems

The RCMM method has been used to evaluate potential energy and force in bulk systems where periodic boundary conditions are used in three dimensions. However, one often encounters systems that are finite in one of the dimensions, such as adsorption and diffusion in slit pores. It is desirable to have a fast and efficient method to treat quasi-2D systems. One of the solutions is to use the EW3D for quasi-2D systems by introducing a large vacuum gap on the top of the slab. The height of the vacuum needs to be adjusted so that it avoids artificial effects in this direction. The method is computationally very slow. Various other approaches have been developed to deal with qausi-2D systems. Parry[9] adapted the Ewald transformation to a laminar and semi-infinite system. Heyes and co-workers derived surface formulas for point–charge and point–dipole[10–16] interactions to calculate potential energy and force in molecular simulations. These two methods are found to be most accurate, but the direct use of these Ewalds 2D formulas is known to be computationally very expensive.[24] Lekner[17] developed a simple surface potential formula cast entirely as a Fourier series. However, this method was devised originally for square systems. Although the scheme has been extended to rectangular systems, its applicability is still limited.[22,23] Hautman and Klein[18] (HK) developed a novel expansion procedure, in which the $1/r$ interaction was decomposed into the in-plane and out-of-plane components in real space. But the HK method was only applied for the system where the distribution of charge ions in the z direction is small.[22]

In this work, a generalized quasi-2D CMM (CMM2D) method was developed by combining the Heyes method with RCMM for confined systems. The methodology of CMM and RCMM described in the previous section still can be applied to quasi-2D system. But, the periodic boundary condition is eliminated in the z direction, and the conventional EW3D summation used in RCMM to deal with fictitious charges was replaced by the Heyes method. For a quasi-2D system, the original unit cell land its 26 neighbor cells are treated by the CMM method to calculate LJ short-range and Columbic long-range interactions. The distant cells with 35 fictitious charges in the RCMM part are treated by the Heyes method to handle long-range interactions. In the Heyes method, the surface lattice is constructed from layers of unit cells infinite in the x and y directions. The 2D real-space lattice vector $\mathbf{n}$ is denoted as $\mathbf{n}=(n_xL,n_yL,L)$ with n_x,n_y integers and reciprocal lattice vector $\mathbf{k}=2\pi\mathbf{n}/L$. The real-space part of the potential energy is shown to be

$$V_{\text{real}}=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{|\mathbf{n}=0|}^{\infty}{}' \; q_iq_j\frac{\operatorname{erfc}(\eta|\mathbf{r}_{ij}+\mathbf{n}|)}{|\mathbf{r}_{ij}+\mathbf{n}|}, \tag{5}$$

where the adjustable parameter η is an arbitrary inverse-length parameter; the $\operatorname{erfc}(x)$ ($\operatorname{erfc}(x)=(2/\sqrt{\pi})\times\int_x^\infty\exp(-t^2)dt$) is complementary error function, which falls to 0 with increasing x. The value of η determines relative emphasis given to the real- and reciprocal-space terms. If η increases, the real-space terms become less important and the reciprocal-space term became more important due to the $\operatorname{erfc}(x)$ function. The prime indicates that the case $i=j$ must be omitted for $\mathbf{n}=0$ since a particle cannot interact with itself. The reciprocal-space contribution to be potential energy is shown to be

$$V_{\text{reciprocal}}=\frac{1}{2}\frac{\pi}{A}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\mathbf{K}} q_iq_jF_{\mathbf{K}}\cos(\mathbf{K}\cdot\mathbf{r}_{ij}), \tag{6}$$

where r_{zij} is the Z component of $\mathbf{r}_{ij}$. The in-plane area of the simulation cell, A is equal to $A=|\mathbf{L}_x\times\mathbf{L}_y|$:

$$F_{\mathbf{K}\neq 0}=\frac{e^{(Kr_{zij})}\operatorname{erfc}\left(\frac{K}{2\eta}+r_{zij}\eta\right)+e^{(-Kr_{zij})}\operatorname{erfc}\left(\frac{K}{2\eta}-r_{zij}+\eta\right)}{K}, \tag{7}$$

$$F_{\mathbf{K}=0}=-2\left\{r_{zij}\operatorname{erf}(r_{zij}\eta)+\frac{e^{-(r_{zij}\eta)^2}}{\eta\sqrt{\pi}}\right\}. \tag{8}$$

The existence of a distinct nonzero term with $\mathbf{K}=0$ is one of the features of the surface formula that distinguishes it from the corresponding bulk Ewald expression,[16] for which the comparable term is equal to zero for an overall charge neutral system. The self-energy term should be subtracted from the total potential energy as for the bulk case,

$$V_{\text{self}}=-\frac{\eta}{\sqrt{\pi}}\sum_{i=1}^{N} q_i^2. \tag{9}$$

The final result for Coulombic interactions in quasi-2D system is

$$V=V_{\text{real}}+V_{\text{reciprocal}}-V_{\text{self}}. \tag{10}$$

Recently, Kawata *et al.*[26] used a Fourier integral for the complimentary error function in Eq. (7) to improve computational efficiency. Minárý *et al.*[28] designed a new formalism in the reciprocal space to treat long-range interactions in quasi-2D systems. This formalism can be implemented in the standard plane-waved density function theory, Ewald, and

the smooth particle-mesh Ewald method for surface calculations. In order to check our CMM2D simulation results, we applied the EW3D technique to quasi-2D systems by adding a large vacuum space on the top of the slab. The inclusion of the vacuum space into the unit cell was done to avoid an artificial influence from periodic images in the z direction.[19–21] Spohr[21] compared the results from this approach with those from the two-dimensional Ewald summation (eW2D) method which was first introduced by Parry[9] and later independently derived by Heyes, Barber, and Clarke.[10] It was concluded that results for this approach converged to those of EW2D when the vacuum height was large in the z direction. By increasing the height of the vacuum space in the z direction, the slab with the vacuum space on its top can be modeled as a strict slab system. The real-space lattice sum decreases quickly because $1/|\mathbf{r}_{ij}+\mathbf{n}|$ falls off very fast in the z direction. The reciprocal vector $\mathbf{K}=2\pi\mathbf{n}/L$ also leads to a decrease in both $\exp(-K^2/4\eta^2)$ and $4\pi^2/K^2$ terms due to the large vacuum in the z direction. Thus, interactions between the original simulation box and its image cells in the z direction with large vacuum height for a bulk system were small. Potential energy will converge within a range of the vacuum height. Jorge and Seaton[27] tested a quasi-2D system with water molecules by using full 2D Ewald sum and 3D Ewald sum methods with large empty spaces in the z direction. Their results show that the potential energy form the 3D Ewald sum converges to the value obtained from 2D Ewald sum with errors below 1%.

C. CMM in GCMC simulations

In grand canonical ensemble MC (GCMC) simulations, chemical potential is fixed while the number of particles fluctuates. The simulations are carried out at constant μ, V, T (chemical potential, volume, and temperature). In GCMC simulations, there are three different types of trials: (a) a molecule is moved, (b) a molecule is destroyed, and (c) a molecule is created at a random position. Each simulation step consists of one of the three attempts described above. For a move attempt, the probability of a movement attempt being accepted is

$$P_{\text{move}}^{\text{acc}}=\min[1,\exp(-\beta\Delta U(r))]. \tag{11}$$

The probability of the creation attempt being accepted is

$$P_{\text{creation}}^{\text{acc}}=\min[1,\exp(-\beta\Delta U(r))+B_i-\ln(N_i+1)]. \tag{12}$$

The probability of deletion being accepted is

$$P_{\text{delete}}^{\text{acc}}=\min[1,\exp(-\beta\Delta U(r))+\ln N_i-B_i]. \tag{13}$$

Here, $\beta=1/k_BT$, $\Delta U(r)$ is the change in confirmation energy; N_i is the number of molecules; V is the volume; and B is defined as

$$B_i=\beta\mu_i+\ln V, \tag{14}$$

where μ_i is the configurational chemical potential of component i.

For creation attempt, when a new molecule is inserted to the original simulation box, it is assigned to a specific cell based on its coordinates. Then, both the molecule coordinates and the cell number for the inserted molecule are labeled. The new trial configuration is updated in the original simulation box and copied to the image cells. It should be pointed out that our CMM update cost is computationally cheap, i.e., $O(1)$ for each trial in GCMC simulations. From the hierarchical tree structure, it is easy to identify near, far and $\infty-27$ reduced cells for the labeled cell. Finally, potential energy between this molecule in the labeled cell and the updated system including the inserted molecule and its image cells is calculated using the methods discussed above. The same procedure is used for deletion and movement attempts. The GCMC-CMM program was tested for a binary system including dendrimer and water. Recently, Jorge and Seaton[27] studied long-range interactions in GCMC simulations of water adsorption in a slit pore using various 2D Ewald sum methods, except for CMM.

III. RESULTS AND DISCUSSION

In this work, applications of the CMM method to MD and GCMC simulations of bulk and confined systems will be discussed. The computational speed and accuracy of the CMM method for evaluating potential energy and force were compared with those from the Ewald summation, minimum image (MI), and Massively Parallel Simulation (MPSim). MPSim[1] is a parallel simulation program with the CMM method developed by Professor Goddard III's group at the California Institute of Technology. The Ewald summation was also available from MPSim. Timing reported for various methods is based on a 400 MHz Silicon Graphics O2 R12K workstation. All structures were built using Cerius[2] from Accelrys, Inc.

A. MD simulations of bulk systems

Tests were first performed on the model of water. The LJ and charge parameters for water are $\sigma_{\text{H}}=2.886$ Å, $\epsilon_{\text{H}}/k=22.144$ K, $q_{\text{H}}=0.41e$, $\sigma_{\text{O}}=3.50$ Å, $\epsilon_{\text{O}}/k=30.196$ K, and $q_{\text{O}}=-0.82e$ from the universal force field (UFF). The system consists of 1664 water molecules with a dimension of 36 Å×36 Å×36 Å. Results are listed in Table I. Tables II and III present the energy decompositions for VDW and Coulombic interactions.

Regarding the level of CMM, as compared with level 1, computational time was reduced by ~6.0 times at level 2 for both the simple and complex downward algorithms, and by 8.8 and 12 times at level 3 for the simple and complex downward algorithms, respectively. Most computational time at level 1 was spent in calculating the near-field interactions the amount to $8\times27\times624^2=84{,}105{,}216$ pairwise interactions. However, at levels 2 and 3, the total near-field pairwise interactions dramatically decreased to $64\times27\times78^2=10{,}513{,}152$ and $512\times27\times10^2=1{,}382{,}400$, respectively. Computational time in the CMM program is the sum of three terms: time for (a) near-field, (b) far-field, and (c) the remaining $\infty-27$ cell interactions. The first term (a) depends on the average number of atoms N in the deepest cell and is proportional to $O(N^2)$. The more atoms are in the deepest cell, the more computational time it will take. Computational time for the second term (b) is approximately linear with $O(K)$, where K is the number of far cells. The last term (c)

TABLE I. Computation of energy, RMS force, and CUP time for 1664 water molecules in the bulk from the Ewald summation, MPSim, minimum image, and our CMM3D program.

Method	Tayor expansion	Level	M[b]	n[c]	Energy[d] (kcal/mol)	RMS[e] (kcal/mol Å)	Time (s)
Ewald					(−13035.37)	(5.166)	668
Min. image					−619.07	0.648	48
MPSim					−50.58	0.013	73
CMM	*Simple*[a]	1	8	624	−58.76	−0.007	210
	Complex[a]						
MPSim					−31.40	0.034	18
CMM	*Simple*[a]	2	64	78	−36.45	0.008	34
CMM	*Complex*[a]				−44.76	0.015	30
MPSim					−4.28	0.032	9
CMM	*Simple*[a]	3	512	9.8	3.33	0.030	24
CMM	*Complex*[a]				−4.63	0.027	17

[a]Simple or complex downward algorithm in the Taylor expansion described in Sec. II and no complex downward for level 1.
[b]Number of cells M in the deepest level.
[c]Average number of atoms n in the leaf cell.
[d]The potential energy is given for the Ewald summation in parentheses. Only the differences are given for other methods.
[e]RMS$=\sqrt{\Sigma_{i=1}^{n} f_i^2/3n-3}$.

is constant at each level for a given system. Thus, there exists an optimal level at which a large number of interactions in the near-field will be compensated by relatively fewer interactions in the far-field. Compared with the Ewald summation method, an error in total energy at level 3 for CMM is less than 0.035%, and an error in the root mean square (RMS) force is about 0.52%, where the percentage errors for energy and RMS are the difference between calculated values from the CMM and Ewald methods divided by that from the Ewald method.

Regarding the simple and complex downward algorithms in Taylor expansion, the latter is faster than the simple downward while maintaining the same order of errors in potential energy and RMS force. In the complex downward procedure, the Taylor series is obtained at the center of a parent cell from distant cells 2 (see Fig. 1). Then, it is shifted to the center of each child cell by multiplying the distance between the centers of the parent cell and its child cell [see Eq. (4)]. Starting at the level-1 cells and recursively repeating this procedure, we can compute the Taylor coefficients $\xi^{(i)}$ for all the cells at all levels. The complex downward expands the Taylor series only once (to the center of the parent cell instead of eight times (to each child cell) in the simple downward. It can be seen in Table I that at level 3, computational time is improved by 29% while accuracy remains similar. At level 2, there is no apparent difference in CPU time and accuracy between the simple and complex downward algorithms since there are a large number of near-field interactions involved. The current program will significantly speed up if data structures (e.g., indexes for each atom and each cell) are optimized.

We also have studied dendrimers, water–methane binary mixtures, as well as water–methane–dendrimer ternary mixtures with various compositions. The PAMAM dendrimer has an ammonia core and $-CH_2CH_2CONHCH_2CH_2NH_2$ monomer units. Additional layers of monomer units can attach to the nitrogen atoms of monomers such that the dendrimer grows like a tree (in contrast to a single chain polymer). PAMAMs of generation 3 (G-3), 4, 5, and 6 have 382, 814, 1678, and 3406 atoms, respectively. Nonbond interactions come from both inter- and intramolecular contributions, excluding 1–2 and 1–3 interactions. PAMAM dendrimers were built using POLYGRAF from Professor William A. Goddard III at the California Institute of Technology. The UFF force field was used to obtain LJ parameters, and the charge equilibration (QEq) method was applied to assign charge to each atom in PAMAM dendrimers. For all LJ cross-term parameters, the geometric mean combining rule was used, namely $\epsilon_{ij}=\sqrt{\epsilon_{ii}\cdot\epsilon_{jj}}$ and $\sigma_{ij}=\sqrt{\sigma_{ii}\cdot\sigma_{jj}}$. For methane, the UFF force field and QEq method were used to obtain LJ parameters and charges. The optimal results for butyl systems are given in Table IV. Results show that the level of CMM (or average atom occupancy) is a key parameter in determining computational time and accuracy. The optimal value for average atom occupancy in the leaf cell is

TABLE II. Energy decompositions for van der Waals and Coulombic interactions at each level for 1664 water molecules in the bulk from our CMM3D program.

Taylor expansion	Level	n	van der Waals sum (kcal/mol)		Electrostatic sum (kcal/mol)	
			Dispersion	Repulsion	CMM	RCMM
Simple	1	624	−8782.65	7004.30	−10478.83	−836.95
Simple	2	78	−8780.57	7004.30	−10458.59	−836.95
Complex	2	78	−8780.46	7004.30	−10467.02	−836.95
Simple	3	9.8	−8762.93	7003.72	−10435.88	−836.95
Complex	3	9.8	−8761.93	7003.72	−10444.84	−836.95

TABLE III. Energy from near- and far-field contributions for van der Waals and Coulombic interactions on each level for 1664 water molecules in the bulk from our CMM3D program.

Taylor expansion	Level	n	van der Waals dispersion (kcal/mol)		van der Waals repulsion (kcal/mol)		Coulomb (kcal/mol)		
			Near	Far	Near	Far	Near	Far	RCMM
Simple	1	624	−8779.39	−3.26	7004.30	1.28E-6	−11660.39	1181.56	−836.95
Simple	2	78	−8751.16	−29.41	7004.30	6.58E-4	−12067.06	1600.04	−836.95
Complex	2	78	−8751.16	−29.30	7004.30	6.58E-4	−12067.06	1600.04	−836.95
Simple	3	9.8	−8523.16	−239.77	7003.38	0.34	−8430.35	−2005.53	−836.95
Complex	3	9.8	−8523.16	−238.77	7003.38	0.34	−8430.35	−2014.48	−826.95

3–10 atoms. Under the optimal condition, the complexdownward algorithm was preferred since it is generally faster than the simple downward method with similar accuracy. Energy error is about 0.05%, while the RMS force error is 0.27%. The RMS force error in polymer systems is much smaller than that in those systems containing water or methane. In CMM, atoms in the same water (or methane) molecule may be divided into different cells. In this case, nonbonded interactions between water molecules are changed from dipole–dipole interactions to charge–dipole or charge– charge interactions. Thus, accuracy will be significantly improved if the atom in the same group whose total charge is zero (e.g., a whole water molecule or each residue in a protein) can be assigned to the same cell as the deepest level.

B. MD simulations of confined systems

In this work, the EW3D as applied to quasi-2D systems was used to check CMM2D results. Figure 2 shows that potential energy and RMS force in dendrimer systems converge within a certain range of vacuum height in the z direction using the EW3D technique.[20,21] According to Spohr[21] and Jorge and Seaton,[27] the conventional EW3D technique with an appropriate vacuum space in the z direction reproduced EW2D results with an energy error of less than 1%, but it is still quite time consuming. Since the EW3D as applied to quasi-2D is our reference and its error in energy as compared to strict EW2D is within 1%, a comparison between our CMM2D and this reference is not as strict as in 3D systems.

A shown in Table V, computational time decreases dramatically as the level increases to $N>1000$. For example, for a G-5 dendrimer having 1678 atoms, as compared with level 1, computational time decreases by 4.0 and 6.6 times at levels 2 and 3, respectively. For a G-6 dendrimer having 3406 atoms, computational time decreases by 4.8 and 8.6 times at levels 2 and 3, respectively. Thus, the larger the system is, the more computationally efficient the CMM method is. For large systems computational time is approximately linear with the number of atoms. At level 1, since there are a large number of atoms in leaf cells, the near-field interactions dominate overall computations. Thus, computational time depends quadratically (N^2) on the number of atoms. The final optimal results for pure components (water or dendrimers), water–methane binary mixtures, and water–methane–

TABLE IV. CPU time, RMS force, and relative energy error for bulk systems with the optimal parameter n and using the complex downward algorithm

	CMM						M.I.		
	N[a]	Level	n[b]	Energy (%)	RMS (%)	Time (s)	Energy (s)	RMS (%)	Time (s)
Pure									
Water	4992	3	9.8	0.04	0.520	17	4.75	12.54	48
dendrimer	382	2	6	0.003	0.054	1	1.91	0.07	1
	814	2	13	0.08	0.080	2	1.14	0.06	1
	1678	3	3.3	0.05	0.017	6	0.28	0.08	4
	3406	3	7	0.04	0.020	28	0.53	0.08	19
Binary mixtures									
water–methane	462	2	7.2	0.04	0.192	1	1.13	6.51	1
	609	2	9.5	0.05	0.583	1	2.84	8.46	1
	1848	3	3.6	0.07	0.630	5	0.85	1.66	6
	3440	3	6.7	0.04	0.514	12	1.65	2.23	26
Ternary mixtures									
dendrimer–water methane	422	2	6.6	0.08	0.0001	1	1.84	1.14	1
	891	2	14	0.06	0.35	2	1.43	0.64	1

[a]Number of atoms in the simulation.
[b]Average number of atoms n in the leaf cell.

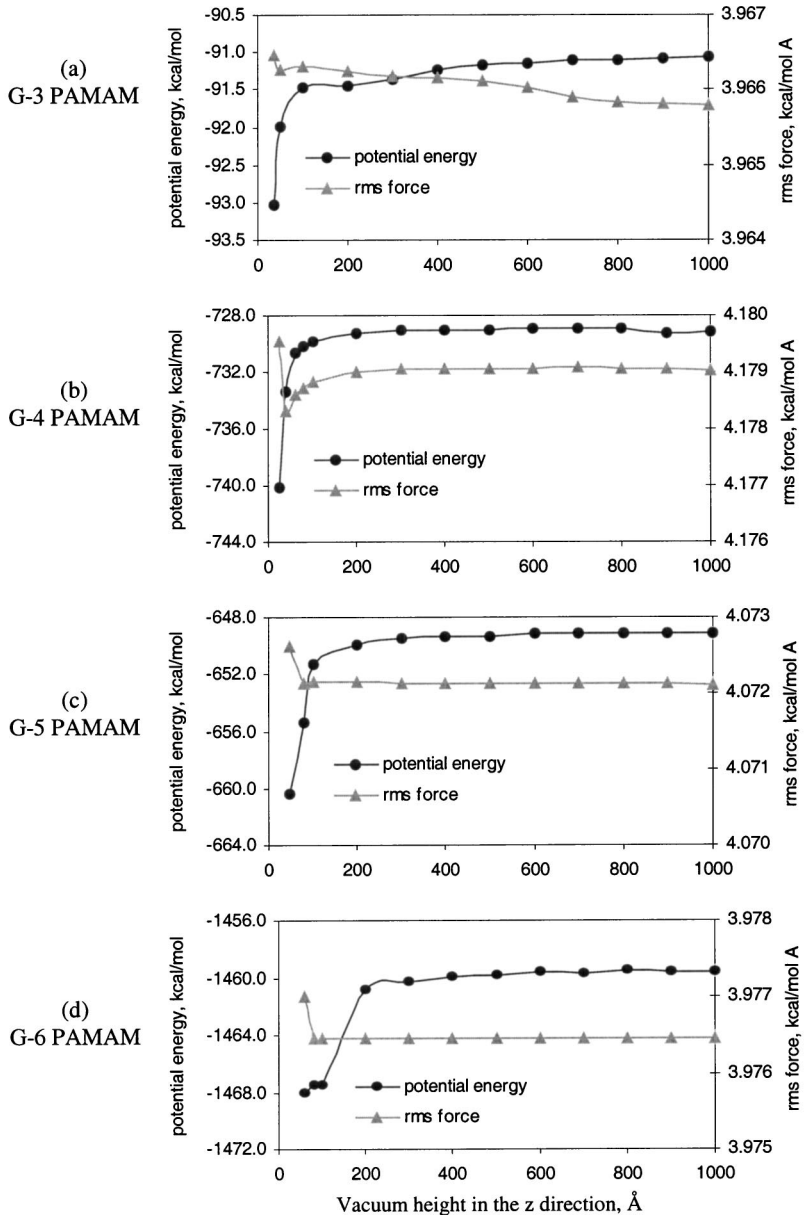

FIG. 2. Potential energy and RMS force vs vacuum height in the z direction for PAMAM dendrimers in the confined systems from the EW3D. (a) $E=-91.083$ and RMS force=3.9658 for G-3 PAMAM with 382 atoms. (b) $E=729.051$ and RMS force=4.1791 for G-4 PAMAM with 814 atoms. (c) $E=649.122$ and RMS force=4.0721 for G-5 PAMAM with 1678 atoms. (d) $E=-1459.610$ and RMS force=3.9765 for G-6 PAMAM with 3406 atoms.

dendrimer ternary systems listed in Table VI, in which the optimal level (or average atom occupancy) and the complex downward algorithm are shown. Comparing the accuracy of CMM with that of the Ewald summation, the error in potential energy at levels 2 and 3 is comparable and is less than 0.32% for various systems. The RMS force error is about 0.04% for dendrimer systems and 0.65% for water or methane systems. In general, CMM is faster than the MI method when the system contains more than 1000 atoms and is much more accurate. The optimal value for average atom occupancy is 3–10 atoms at the deepest level. The complex downward algorithm is generally faster than the simple one with similar accuracy.

C. GCMC simulations of confined systems

A system with a dimension of 46.7 Å×59.6 Å×57.3 Å was tested for a binary mixture of a G-5 PAMAM dendrimer and water in a confined system using our CMM-GCMC program. The difference in energy ΔE between two different configurations was calculated at each level for each GCMC move. Results in Table VII show that potential energies calculated from our CMM-GCMC program for various GCMC moves are consistent with those from our CMM-MD program. The slight difference among different levels for each move is due to approximations used in the calculation of far-field contributions.

IV. CONCLUSIONS

The CMM method is efficient in calculating both van der Waals and Coulombic interactions with significantly low computational time and good accuracy. It is well suited to handling large systems with long-range interactions in molecular simulations. Simulation results show that the level of the cell hierarchical tree (or average atom occupancy) in the deepest cell is a key factor in determining computational time and accuracy. The optimal number of atoms in the deep-

TABLE V. Computation of energy and CPU time for one G-3, G-4, G-5, or G-6 PAMAM dendrimer in the confined system from the Ewald sum and our CMM2D program.

N	Method	Taylor expansion	Level	*M*	*n*	Energy (kcal/mol)	RMS (kcal/mol Å)	Time (s)
382	Ewald					(−91.083)	(3.9658)	
	M.I.					−0.935	0.0040	1
	CMM	*Simple*	1	8	48	−1.803	0.0020	1
	CMM	*Simple*				−1.935	0.0015	1
			2	64	6			
	CMM	*Complex*				−0.067	0.0004	1
	CMM	*Simple*				0.406	0.0008	5
			3	512	0.7			
	CMM	*Complex*				−1.719	0.0026	4
814	Ewald					(−729.051)	(4.1791)	
	M.I.					−3.831	0.0035	1
	CMM	*Simple*	1	8	102	−3.754	0.0024	4
	CMM	*Simple*				−3.748	0.0023	1
			2	64	12.7			
	CMM	*Complex*				−1.652	0.0024	1
	CMM	*Simple*				6.859	0.0103	5
			3	512	1.6			
	CMM	*Complex*				5.558	0.0103	4
1678	Ewald					(−649.122)	(4.0721)	
	M.I.					−6.122	0.0035	4
	CMM	*Simple*	1	8	210	−2.420	0.0012	32
	CMM	*Simple*				7.393	0.0052	9
			2	64	26			
	CMM	*Complex*				7.630	0.0059	8
	CMM	*Simple*				5.113	0.0007	6
			3	512	3.3			
	CMM	*Complex*				6.008	0.0007	4
3406	Ewald					(−1459.610)	(3.9765)	
	M.I.					−9.872	0.0002	17
	CMM	*Simple*	1	8	428	5.916	0.0019	172
	CMM	*Simple*				−768	0.0014	36
			2	64	63			
	CMM	*Complex*				−3.474	0.0014	36
	CMM	*Simple*				−8.607	0.0008	21
			3	512	6.7			
	CMM	*Complex*				−1.152	0.0005	19

TABLE VI. CPU time and percentage error in potential energy with an optimum level and using the complex downward algorithm for pure (water or PAMAM dendrimeter), binary mixture (water and methane), and ternary mixture (water, methane, and dendrimer) in the confined system from our CMM2D program.

	CMM						M.I.		
	N^a	Level	n^b	Energy (%)	RMS (%)	Time (s)	Energy (s)	RMS (%)	Time (s)
Pure									
Water	4992	3	9.8	0.21	0.04	12	0.89	0.08	42
dendrimer	382	2	6	0.07	0.01	1	1.03	0.10	1
	814	2	13	0.22	0.06	1	0.53	0.08	1
	1678	3	3.3	0.92	0.08	4	0.94	0.09	4
	3406	3	7	0.08	0.01	19	0.67	0.005	17
Binary mixtures									
water–methane	462	2	7.2	0.22	1.32	1	2.11	8.32	1
	609	2	9.5	0.44	.82	1	2.53	5.94	1
	1848	3	3.6	0.15	0.89	3	0.20	1.62	5
	3440	3	6.7	0.12	0.58	8	0.29	1.32	24
Ternary mixtures									
dendrimer–water methane	422	2	6.6	0.86	0.81	1	3.74	1.51	1
	891	2	14	0.28	0.09	1	1.21	0.10	1

TABLE VII. A comparison of potential energy from our CMM-MD and CMM-GCMC programs.

		CMM-MD			CMM-GCMC
Add		Old (kcal/mol)	New (kcal/mol)	ΔE (kcal/mol)	ΔE (kcal/mol)
Level 1	*Simple*	242.18	242.108	−0.072	−0.068
Level 2	*Simple*	232.783	232.791	0.008	0.004
	Complex	231.28	231.293	0.013	0.017
Level 3	*Simple*	233.733	233.717	−0.016	−0.025
	Complex	232.683	232.712	0.029	0.033
		CMM-MD			CMM-GCMC
Delete		Old (kcal/mol)	New (kcal/mol)	ΔE (kcal/mol)	ΔE (kcal/mol)
Level 1	*Simple*	242.108	242.181	0.072	0.074
Level 2	*Simple*	232.791	232.783	−0.008	0.0003
	Complex	231.293	231.28	−0.013	−0.012
Level 3	*Simple*	233.717	233.733	0.016	0.029
	Complex	232.712	232.683	−0.029	−0.028
		CMM-MD			CMM-GCMC
Move		Old (kcal/mol)	New (kcal/mol)	ΔE (kcal/mol)	ΔE (kcal/mol)
Level 1	*Simple*	242.1079	242.4757	0.368	0.364
Level 2	*Simple*	232.7917	233.2285	0.437	0.433
	Complex	231.2931	231.7315	0.438	0.435
Level 3	*Simple*	233.7171	234.1922	0.475	0.455
	Complex	232.7122	233.211	0.499	0.495

est cell is 3–10. The complex Taylor expansion algorithm is faster than the simple one with similar accuracy. Under the optimal conditions, the average error in total potential energy is about 0.05% and the RMS force is 0.27% when compared with the Ewald summation method while computational efficiency improves significantly. Results show that the computational time of the CMM method scales almost linearly with the number of atoms in the cell for large systems.

The CMM method was extended to confined systems in MD simulations. Potential energies calculated from our CMM2D program are consistent with those from the EW3D technique applied to quasi-2D systems by adding a sufficiently large vacuum space on the top of the slab. The difference in total potential energy between these two methods is 0.32% and RMS force error 0.43%. Moreover, the CMM method was applied to GCMC simulations for both 3D and 2D systems. Our CMM code is readily incorporated in the GCMC, MD, and grand canonical molecule simulation (GCMD) programs for applications, such as studies of adsorption and diffusion in pores.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Department of Energy and the University of Washington for financial support. The facilities of the MSC are supported by grants from DoE ASCI ASAP, ARO-MURI, BP Amoco, Chevron Corp., NASA, Beckman Institute, Seiko-Epson, Exxon, Asahi Chemical, Avery-Dennison, Dow, and 3M.

[1] K.-T. Lim *et al.*, J. Comput. Chem. **18**, 501 (1997).
[2] H.-Q. Ding, N. Karaawa, and W. A. Goddard III, J. Chem. Phys. **97**, 4309 (1992).
[3] H.-Q. Ding, N. Karasawa, and W. A. Goddard III, Chem. Phys. Lett. **196**, 6 (1992).
[4] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[5] A. Y. Toukmaji and J. A. Board, Jr., Comput. Phys. Commun. **95**, 73 (1996).
[6] A. W. Appel, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **6**, 85 (1985).
[7] J. Carrier, L. Greengard, and V. Rokhlin, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **9**, 669 (1989).
[8] P. Vashishta *et al.*, Mater. Sci. Eng., B **37**, 56 (1996).
[9] D. E. Parry, Surf. Sci. **49**, 433 (1975); **54**, 195 (1976).
[10] D. M. Heyes, M. Barber, and J. H. R. Clarke, J. Chem. Soc., Faraday Trans. 2 **73**, 1485 (1977).
[11] D. M. Heyes, J. Phys. Chem. Solids **41**, 291 (1980).
[12] D. M. Heyes and F. van Swol, J. Chem. Phys. **75**, 5051 (1981).
[13] D. M. Heyes, J. Chem. Phys. **79**, 4010 (1983).
[14] D. M. Heyes, Phys. Rev. B **30**, 2182 (1984).
[15] D. M. Heyes, Surf. Sci. **110**, 1619 (1981).
[16] D. M. Heyes, Condens. Matter News **49**, 755 (1994).
[17] J. Lekner, Physica A **176**, 485 (1991).
[18] J. Hautman and M. L. Klein, Mol. Phys. **75**, 379 (1992).
[19] S. W. de Leeuw, J. W. Perram, and E. R. Smith, Proc. R. Soc. London, Ser. A **373**, 27 (1980).
[20] S. Jiang, R. Frazier, E. S. Yamaguchi, M. Blanco, S. Dasgupta, Y. Zhou, T. Cagin, Y. Tang, and W. A. Goddard III, J. Phys. Chem. B **101**, 7702 (1997).
[21] E. Spohr, J. Chem. Phys. **107**, 6342 (1997).
[22] S. Y. Liem and J. H. R. Clarke, Mol. Phys. **92**, 19 (1997).
[23] A. Grzybowski, E. Gwóźdź, and A. Bródka, Phys. Rev. B **61**, 6706 (2000).
[24] In-Ch. Yeh and M. L. Berkowitz, J. Chem. Phys. **111**, 3155 (1999).
[25] W. A. Goddard III, MSC Technical Notes, Material and Molecular Simulatin Center at Caltech, 1994
[26] M. Mawata and M. Mikami, Chem. Phys. Lett. **340**, 157 (2001).
[27] M. Jorge and N A. Seaton, Mol. Phys. **100**, 2017 (2002).
[28] P. Minárý, M. W. Tuckerman, K. A. Pihakari, and G. J. Martyna, J. Chem. Phys. **116**, 5351 (2002).