Journal of Computational Physics 270 (2014) 613–633

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

An improved fast multipole method for electrostatic potential calculations in a class of coarse-grained molecular simulations

Mohammad Poursina [a], Kurt S. Anderson [b]

[a] Department of Aerospace and Mechanical Engineering, University of Arizona, Tucson, AZ 85721, United States
[b] Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, United States

ARTICLE INFO

Article history:
Received 13 November 2013
Received in revised form 1 April 2014
Accepted 8 April 2014
Available online 16 April 2014

Keywords:
Fast multipole method
Multibody dynamics
Far-field interactions
Coarse-graining
Molecular dynamics
Electrostatic potential field

ABSTRACT

This paper presents a novel algorithm to approximate the long-range electrostatic potential field in the Cartesian coordinates applicable to 3*D* coarse-grained simulations of biopolymers. In such models, coarse-grained clusters are formed via treating groups of atoms as rigid and/or flexible bodies connected together via kinematic joints. Therefore, multibody dynamic techniques are used to form and solve the equations of motion of such coarse-grained systems. In this article, the approximations for the potential fields due to the interaction between a highly negatively/positively charged pseudo-atom and charged particles, as well as the interaction between clusters of charged particles, are presented. These approximations are expressed in terms of physical and geometrical properties of the bodies such as the entire charge, the location of the center of charge, and the pseudo-inertia tensor about the center of charge of the clusters. Further, a novel substructuring scheme is introduced to implement the presented far-field potential evaluations in a binary tree framework as opposed to the existing quadtree and octree strategies of implementing fast multipole method. Using the presented Lagrangian grids, the electrostatic potential is recursively calculated via sweeping two passes: assembly and disassembly. In the assembly pass, adjacent charged bodies are combined together to form new clusters. Then, the potential field of each cluster due to its interaction with faraway resulting clusters is recursively calculated in the disassembly pass. The method is highly compatible with multibody dynamic schemes to model coarse-grained biopolymers. Since the proposed method takes advantage of constant physical and geometrical properties of rigid clusters, improvement in the overall computational cost is observed comparing to the tradition application of fast multipole method.

1. Introduction

Modeling and simulation of biopolymers such as proteins, DNAs, and RNAs can provide significant information about the important biological phenomena which occur inside living cells. However, high frequency motions of atoms impose a constraint on the size of the integration time step of explicit integrators (0.5–2 fs) [1,2] when such systems are modeled using the fully atomistic representation. This is undesirable since many biologically important processes occur on time scales as slow as milliseconds to seconds [3]. Furthermore, atomistic resolution models of such complex systems with large number of particles ($n \approx 10^6$ [4]) suffer from the cumbersome pairwise force field determinations with computational complexity of $O(n^2)$ at each time step.

To facilitate long-time molecular simulations, different techniques have been developed to either increase the size of the integration time step, or reduce the computational cost per integration time step, or the combination of these approaches.

http://dx.doi.org/10.1016/j.jcp.2014.04.025
0021-9991/

These improvements may be achieved through the intelligent use of specialized coarse-grained models via removing high frequency modes of motion of the system. Such models which facilitate the application of the integration time step beyond 20 fs [1,2,5] may be obtained by assuming groups of atoms as spherical pseudo-atoms or beads [6–10]. As such, particle dynamics is used to form the equations of motion of larger rigid pseudo-atoms. Coarse graining may also be implemented by considering groups of atoms as rigid and/or flexible bodies connected together via kinematic joints. These kinematic joints through which one can capture the gross motion of the system are able to represent the translational and/or rotational motion between the bodies. In this scheme which has been recently used in a variety of applications such as modeling polymers such as DNAs, RNAs, and proteins [2,11–17], geometric constraints such as fixing bond lengths can be exactly enforced. Moreover, the mass distribution and geometry of each articulated pseudo-atom as well as the resultant torques applied to these clusters are included to form and solve the equations of motion. It is shown in [18] that ignoring torques to address the rotational motion of the clusters, which is essentially the case in bead models, may end up with unrealistic configurations.

The dominant portion of the computational cost of molecular simulations which usually accounts for over 90% of the execution time [19] is associated with evaluating pairwise potential fields. Coarse-grained models reduce these costs to some extent due to ignoring the interactions among the atoms embedded in the same rigid clusters. Moreover, a variety of advanced algorithms has been developed for the efficient far-field determinations. This starts with the intuitive approach of ignoring the interactions beyond the cut-off distance. However, this strategy leads to the artifacts including the poor reproduction of properties of the system [20–23]. Particle–mesh (PM) and particle–particle particle–mesh (PPPM) techniques suggested for non-periodic systems perform satisfactorily when the required precision is low, and particles are distributed more or less uniformly [24–27]. Ewald summation which computes the long-range interactions in Fourier space is only applicable to periodic systems [28,29]. More advanced methods have been developed based on the power series expansions. In this scheme, the monopole approximation of the long-range field in the system is presented in [24,30,31]. Using Fast Multipole Method (FMM) [32] with the application of higher order multipole and Taylor's series expansions, one can find the electrostatic field of a group of particles with a computational complexity of $O(n)$. The FMM method is later extended and used in CPU-based serial and parallel implementations to approximate different potential fields for different systems [3,4,33–39]. Significant computational achievements have also been gained by applying GPUs to molecular modeling for the $O(n^2)$ and cutoff-based $O(n)$ nonbonded force evaluation [40–45] due to their much higher computational efficiency comparing to highly optimized CPU-based codes. The application of GPU to perform $O(n)$ FMM potential evaluations has also been reported in [46].

The fast multipole method is implemented by recursively subdividing the space to cubical cells [24,32]. The number and location of particles in these Eulerian grids (cubical cells) are not fixed. Therefore, the associated multipole and Taylor's series expansion coefficients vary with time. The Eulerian grid scheme works well for problems in which the particles are distributed more or less uniformly over a domain. Although the application of the adaptive fast multipole methods [37,38] improves the potential determinations of the systems which are not uniformly distributed in the space, the corresponding substructuring strategy does not lend itself well to the multibody framework which has been recently used to model chain polymers [2,12–17]. For instance, the article [36] is the only reference (to the best of the authors' knowledge) which applies the Cell Multipole Method (CMM) [4] to evaluate pairwise forces in modeling articulated coarse-grained biopolymers. In this scheme which uses an adaptive octree framework, the chain molecule is placed in a box which is divided into eight cells. The resulting cells are then divided into eight cells as shown in Fig. 1, until there are about four particles in the smallest cell, i.e. "microcell" [4]. Forces applied to the particles are then calculated hierarchically by recursively evaluating the multipole and Taylor's series expansion coefficients associated with cells under the interactions. Based on the schematics shown in this figure, it is observed that some of these cells (grids) contain a portion of a single cluster, while others may contain pieces of different clusters. As such, the location of each particle must be reconciled with the spatial grid at each sampling. This imposes an additional computational burden on the simulation, particularly when the coarse-grained model is used for the simulation purposes. In other words, in addition to tracking the location of the mass center of each rigid cluster of the system as a result of solving the equations of motion, the spatial location of each single particle must also be tracked and reconciled with the spatial grid to update the terms which appear in the force field approximations. Further, this method provides the potential field and/or the force at each particle which is unnecessary in the multibody-based coarse graining approach in which the resultant load applied to the body or the related potential field which is computationally less expensive is sufficient to accurately form and solve the equations of motion. Finally, the available advanced methods do not efficiently benefit from the rigidity of pseudo-atoms of the coarse grained system model which results in constant physical and geometrical properties of these clusters.

In the work presented in [18], the approximate resultant force and moment due to pairwise interactions in the form of $\frac{1}{r^s}$ ($s > 1$ is an integer) between the atoms embedded in a pseudo-atom cluster in the coarse-scale region and a particle which resides in the fine-scale region are calculated. Moreover, that article presents the far-field force and torque due to interactions between the particles embedded in two different clusters in the coarse-scale regions of the system. These approximations are expressed in terms of physical and geometrical properties of the pseudo-atoms which are constant for rigid clusters. However, that article does not provide an efficient and practical mathematical framework to use and implement the developed formulas to find the forces and moments on the bodies.

In this article, a novel recursive algorithm is developed and implemented to efficiently approximate the long-range electrostatic potential field of systems containing highly negatively/positively rigid or flexible clusters. Systems with such