

# Efficient Coarse-Grained Molecular Simulations in the Multibody Dynamics Scheme

Mohammad Poursina and Kurt S. Anderson

**Abstract** The numerical simulation of highly complex biomolecular systems such as DNAs, RNAs, and proteins become intractable as the size and fidelity of these systems increase. Herein, efficient techniques to accelerate multibody-based coarse-grained simulations of such systems are presented. First, an adaptive coarse-graining framework is explained which is capable of determining when and where the system model needs to change to achieve an optimal combination of speed and accuracy. The metrics to guide these on-the-fly instantaneous model adjustments and the issues associated with post-transition system's states are addressed in this book chapter. Due to its highly modular and parallel nature, the Generalized Divide-and-Conquer Algorithm (GDCA) forms the bases for a suite of dynamics simulation tools used in this work. For completeness, the fundamental aspects of the GDCA are presented herein. Finally, a novel method for the efficient and accurate approximation of far-field force and moment terms are developed. This aspect is key to the success of any large molecular simulation since more than 90 % of the computational load in such simulations is associated with pairwise force calculations. The presented approximations are efficient, accurate, and highly compatible with multibody-based coarse-grained models.

## 1 Introduction

Development and application of the efficient techniques to model, simulate, and analyze highly complex biomolecular systems such as DNAs (Deoxyribonucleic acid), RNAs (Ribonucleic acid), enzymes, and proteins have been gaining attention by scientists and engineers in an effort to predict and understand different structural, mechanical, and thermodynamic properties of such systems [17, 22, 44, 56, 71].

---

M. Poursina (✉) · K.S. Anderson

Computational Dynamics Laboratory (CDL), Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute (RPI), Troy, NY 12180, USA  
e-mail: [poursm2@rpi.edu](mailto:poursm2@rpi.edu)

K.S. Anderson

e-mail: [anderk5@rpi.edu](mailto:anderk5@rpi.edu)

These molecular simulations provide important information about the relationships between the structure and function of biopolymers, and provide insight into various biological processes. Fully atomistic representation of such systems [33, 43] results in detailed information on the underlying physics of these biopolymers and captures all scales of the problem. However, these simulations are greatly limited because they cannot be accomplished in a timely manner for models possessing the desired size and fidelity. This is due to the existence of high frequency motions which impose a tight constraint on the temporal integration step size (0.5–2 fs) of explicit integrators [19, 75], while biologically important processes occur on time scales as slow as milliseconds to seconds [55]. Furthermore, these models with large number of atoms ( $n_a \approx 10^6$  [23]) suffer from the cumbersome pairwise force field calculations with the computational complexity of  $O(n_a^2)$  at each time step. As such, different methods have been developed to improve the temporal integration step size and reduce the computational cost per integration time step.

In biomolecular systems, high frequency and low amplitude motions of the atoms are responsible for local motions, while low frequency and high amplitude motions of the system's subdomains are dominant in representing the global conformation. Hence, the simulation performance may be improved significantly via the intelligent use of specialized coarse-grained models in which high frequency modes of motion of the system are removed from consideration. These models still capture the overall conformational motion of the biopolymers while allowing the application of larger integration time steps (e.g., 20–50 fs [19, 31, 75]).

The coarse-grained model may be realized by treating a group of atoms as a spherical bead (pseudo-atom) [16, 31, 66]. For instance, each nucleotide in a protein chain may be modeled using one to six beads [74]. Other applications of such spherical beads in dissipative particle dynamics and solvent lipid interactions are reported in [48, 73].

Alternatively, a group of atoms may be represented by a rigid or flexible body connected to its parent and child bodies via kinematic joints. Using internal coordinates (i.e., generalized coordinates which describe the relative motions of the child bodies with respect to the parent bodies at the connecting joints), the geometric constraints such as fixing bond lengths can be enforced exactly. In the finest coarse-grained articulated body model, dihedral angles (torsion dynamics) are used to describe the dynamics of the system [1, 36, 40, 72], while the bond stretch and bond angles are frozen. Unlike the spherical beads, the mass distribution and geometry of each articulated pseudo-atom is expressed in terms of the associated inertia tensor, and the distance from the corresponding mass center to the joints of the pseudo-atom. In such models, which may be particularly applicable to simulate the dynamic behavior of polymer chains, both translational and rotational motion [11, 18] of each cluster [36] are considered in forming the equations of motion. As such, the effect of Coriolis and centrifugal inertial forcing terms are considered in the equations of motion. As the length of the polymer chain increases, the role of these terms becomes more important in capturing the dynamics of the system due to scaling effects. Since these articulated models address the rotational motion of the rigid and flexible subdomains of the system in the equations of motion, they

can capture the geometry of biopolymers more accurately than most of the bead-based coarse-grained models which ignore the rotational dynamics of the spherical pseudo-atoms.

It is demonstrated in [57, 63] that the dynamic behavior of biopolymers is highly nonlinear, and significantly affected by the change in the kinematics and dynamics of the boundary conditions of the system. As such, static (time-invariant) coarse-grained models may not appropriately capture the dynamics of the system for the entire course of the simulation. This requires the development of adaptive machinery to perform such simulations, particularly when the non-equilibrium behavior of these systems is of interest.

In the adaptive multiscale strategy presented here, some degrees of freedom (internal coordinates) have their definitions/meanings adjusted “on-the-fly” at different instants and different locations of the system based on the values of knowledge-, math-, and/or physics-based metrics. Herein, the appropriate metrics to guide these model transitions are investigated. Each model adjustment towards the lower or higher fidelity system model may be viewed as the instantaneous application or release of system’s internal constraints. As such, the generalized momentum of the system must be conserved to arrive at the physically meaningful post-transition system’s states. It is also demonstrated that within the transitions to the finer-scale models, some issues arise which are associated with the proper amount and placement of the energy within the system.

Given the central role that multibody dynamics plays in the presented framework, a suite of Generalized Divide-and-Conquer Algorithm-based approaches is employed to this end. These methods offer a good combination of computational efficiency and modular structure. Furthermore, the computational complexity of the algorithm is  $O(n)$  and  $O(\log n)$  in serial and parallel implementations, respectively, where  $n$  denotes the number of degrees of freedom of the system.

A key aspect of this book chapter is associated with pairwise force calculations in molecular simulations. More than 90 % of the computational cost per temporal integration step in modeling biopolymers is associated with calculating these forcing terms. Focusing on long-range (far-field) interactions in the system, various methods such as Barnes-Hut [5, 10], Edwald summation [21, 24], and the Fast Multipole Method [32] have been developed to reduce the cost of these forcing terms evaluations. A review on these methods is presented in [57]. Herein, a novel approximation for the far-field force and moment calculations that is well suited for use in an articulated body modeling of biopolymers is presented. This technique may be viewed as a generalization and extension of the method presented in [38] to approximate the gravitational force used in modeling spacecraft dynamics. The resultant force and moment due to the pairwise interactions are approximated between: the atoms (particles) embedded in a pseudo-atom (body) in the coarse-scale region and an atom (particle) which resides in the fine-scale domain; and between the particles embedded in two different pseudo-atoms (bodies) in the coarse-scale regions of the system. The value of the moment ignored in bead-based models can produce significant errors which is more accurately captured using multibody-based coarse-grained simulations. These low order multipole approximations and Taylor

expansions are expressed in terms of the geometric and physical properties of the pseudo-atoms (subdomains). Since these properties are constant for rigid domains of the system, these approximations significantly accelerate the evaluation of the far-field forcing terms.

## 2 The Need for Adaptive Simulation of Molecular Systems

The dynamics of biopolymers is highly nonlinear and chaotic. Although these systems have highly chaotic components, their coarser scale conformational behavior may tend towards a specific structure. As such, it may be possible to remove high frequency modes of motion which contribute little to the final structure, while maintaining more important lower frequency components. However, given the complex nature of the system behavior, it may not be possible to identify which modes of motion can be removed a priori. For instance, the simulations conducted on articulated RNAs with various sequences in [57, 63] show that the dynamic behavior of each joint angle is highly time variant, and is significantly affected by the changes in the dynamics of the rest of the system. Those results demonstrate that the current coarse-grained model which is potentially correct for a specific time interval may not provide accurate and reliable information about the conformational motion of the system for the entire simulation. This results in the inadequacy of the static coarse graining in which the system model does not change within the course of the simulation. As such, adaptive multiscale methods which perform the coarse graining in time and space must be developed to better model the dynamics of biopolymers.

The adaptive machinery is capable of identifying the critical locations of the system to remove and/or add fidelity from/to the system model as necessary. In this process, some degrees of freedom of the system are adaptively constrained or released at different instants and different locations of the system. As such, this framework automatically adjusts the coarseness of the model, in an effort to more optimally increase the simulation speed, while maintaining accuracy. In the following, the necessary machinery to implement the adaptive multiscale simulation of biopolymers is presented. Development of the metrics to guide the model adjustments, efficiently modeling the forward dynamics, as well as appropriately handling the dynamics of the transitions are important aspects in this scheme. More detail on the adaptive multiscale framework to model biopolymers is found in [13, 57, 63].

## 3 Metrics to Steer Transitions

In situations where modification of the dynamics model is appropriate, these adjustments should be guided by suitable internal metrics. These metrics may be physics-based (derived directly from physical laws), knowledge-based (derived empirically) and/or math-based (derived from strictly mathematical relations). Herein, two different metrics which respectively guide the model transitions to the coarser and finer models are investigated.

### 3.1 Fine-to-Coarse Transitions

The transition to the coarser-scale model can be achieved via removing the less significant modes of motion of the system. This can be realized by monitoring the behavior of the individual internal coordinates of the system model to assess which of them may be removed, while not adversely changing the conformational behavior of the system. In molecular simulations, high frequency modes of motion do not contribute significantly to the global conformation of the system, while providing large instantaneous relative velocities and accelerations. As such, velocity- and acceleration-based metrics are not well-suited for identifying the more significant degrees of freedom of the system in the overall conformational motion. Furthermore, these complex systems are highly nonlinear and chaotic; therefore, the instantaneous values of the states of the system are not expected to (and have been shown not to [63]) work well for guiding model transitions.

Monitoring the moving-window statistical properties of the internal coordinates of the system is proposed in [63] as a math-based metric to assess and guide the coarse graining process instead of the instantaneous velocity- and acceleration-based measures described in [67, 70]. The standard deviation of the generalized coordinates defined at the joints (internal coordinates) collected within the moving window as given by

$$S_w = \sqrt{\frac{\sum_{k=1}^{n_w} (q_k - \bar{q}_w)^2}{n_w}} \quad (1)$$

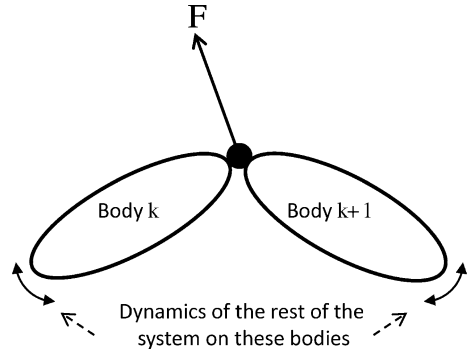
is suggested as the metric of choice to determine if an existing joint should be kept or removed. In this relation,  $\bar{q}_w$  is the moving-window average of the sequence of data  $\{q_k\}_{k=1}^{n_w}$  within the window of the size of  $n_w$ .

It should be mentioned that the overall conformation of the system may be more sensitive to some specific internal coordinates, while this sensitivity varies with time [13, 57]. As such, if the scaled (weighted) moving-window standard deviation of any internal coordinate of the system is less than a predefined threshold, then the associated degree of freedom is eligible to be frozen.

### 3.2 Coarse-to-Fine Transitions

Since the static coarse-grained models may not appropriately predict the overall conformational motion of the biopolymer, it may be necessary to add fidelity to the system model within the course of the simulation. The system's constraints and internal loads as shown in Fig. 1 arise from the kinematic constraints imposed on adjacent body-to-body motions by the connecting joints, the interactions between the bodies, and the imposed boundary conditions. Furthermore, the constraint load indicates the degree to which the body or joint in question is attempting to be deformed at its location. Therefore, monitoring the spatial constraint loads (forces and torques)

**Fig. 1** The constraint load  $F$  gives an estimate of the error introduced because of locking the joint



acting on all kinematic joints of the system, and intermediate locations within the rigid and flexible bodies is the proposed metric to assess the validity of the selected coarse-grained model, and to guide the model refinement. In other words, the joint is released (or added) if the spatial constraint force at the associated location exceeds the nominal load which figuratively causes a mechanical failure [63].

## 4 Generalized Divide-and-Conquer Based Adaptive Framework

The simulation of the biopolymer in an adaptive framework should appropriately and efficiently address the forward dynamics, model adjustments, and dynamics of the transitions. In all of these steps, multibody dynamics plays an important role since herein the pseudo-atoms may be viewed as rigid and/or flexible domains of the system connected together via kinematic joints. As the complexity of these multibody systems (manifesting itself in the form of modes of motion) increases, a prohibitive computational burden can be imposed on the simulation due to the kinematic coupling which exists in most articulated multibody formulations. Different so-called  $O(n)$  algorithms (as opposed to traditional  $O(n^3)$  algorithms) in which the simulation turnaround time scales linearly with the increase in the system's degrees of freedom (i.e.,  $n$ ) have been developed in [2, 6, 8, 15, 25, 26, 35, 45, 54, 68, 69, 76, 77] as an effort to reduce this undesirable scaling in computational effort with problem size. These  $O(n)$  methods are less costly for large  $n$ , but due to their underlying serial recursive under-pinnings, they generally do not lend themselves well to massive parallelization. As such, different algorithms have been designed to model and simulate multibody systems which better exploit the parallel computing capability [7, 9, 20, 29, 30, 34, 39, 42].

The Divide-and-Conquer Algorithm (DCA) is a recursive method of modeling multibody systems first demonstrated by Featherstone [27, 28]. Its recursive structure is not serial, but that of a binary tree. Different extensions and applications of this method in modeling and sensitivity analysis of the multi-rigid/flexible-body systems, studying the impulsive behavior and contact problems in such systems, as well as the parallel implementation of this algorithm are reported in [12–14, 46, 47, 49–53]. In this scheme, a complete set of both absolute and generalized coordinates

are used to form the equations of motion. As such, the analyst can select a set of appropriate coordinates to integrate and work with. Applying both spatial and generalized coordinates, and directly imposing the constraints describing the kinematics of the joints to form the equations of motion of the system, the DCA provides a robust framework to address the dynamics of the kinematically closed-loop systems in singular configurations [50]. A Generalized Divide-and-Conquer Algorithm (GDCA) developed in [57, 60] is an extension to the DCA which can be easily used to model multibody systems in which a part of the forcing information is provided in a generalized force format. This may occur in modeling the systems in which a set of known/unknown generalized forces must be considered in the equations of motion due to the application of the control law or the imposition of the algebraic constraints [58, 59]. For instance, this method is used in [57, 58] to perform the constant temperature simulation of biopolymers in which the feedback forces from the thermostat are provided in the generalized format.

Herein, the GDCA-based methods are suggested to be used in the context of the large-scale adaptive molecular problems because:

1. They are relatively efficient for the large-scale sequential computer implementation. The computational complexity of these algorithms for unconstrained systems is  $O(n)$  in the serial implementation.
2. These methods are highly parallelizable which provide a time optimal order  $\log n$  computational performance achieved with a processor optimal order  $n$  processors.
3. The implementation and use of these formulations within an adaptive framework are relatively straightforward due to the algorithm's highly modular structure.

## 4.1 Forward Dynamics

Similar to the DCA, the dynamics of each body in the GDCA scheme is expressed in terms of the handle equations of motion [50]. A handle is a point of the body through which it has an interaction with its surroundings. Herein, the handle equations in the GDCA scheme are presented to illustrate how they accommodate generalized forces.

Consider an arbitrary body  $k$  shown in Fig. 2 connected to bodies  $k - 1$  and  $k + 1$  via kinematic joints  $J^k$  and  $J^{k+1}$ , respectively. Each degree of freedom of this system is defined as the relative motion of the child body with respect to its parent body. Let the column matrix  $\hat{f}$  contain the known/unknown generalized forces associated with some specific degrees of freedom which must be included in the equations of motion. The two-handle equations of motion for body  $k$  in the GDCA scheme are presented by the following relations [57, 60]

$$\mathcal{A}_1^k = \phi_{11}^k(\mathcal{F}_{1c}^k + \mathcal{D}_s^{J^k} \hat{f}_k) + \phi_{12}^k(\mathcal{F}_{2c}^k - \mathcal{D}_s^{J^{k+1}} \hat{f}_{k+1}) + \phi_{13}^k, \quad (2)$$

$$\mathcal{A}_2^k = \phi_{21}^k(\mathcal{F}_{1c}^k + \mathcal{D}_s^{J^k} \hat{f}_k) + \phi_{22}^k(\mathcal{F}_{2c}^k - \mathcal{D}_s^{J^{k+1}} \hat{f}_{k+1}) + \phi_{23}^k. \quad (3)$$

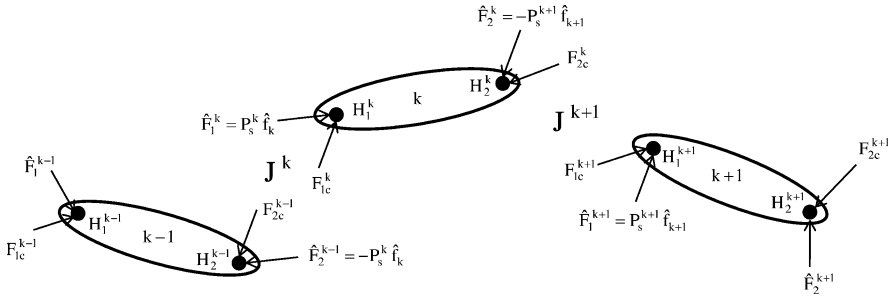


Fig. 2 Interactions between consecutive bodies in a multibody system

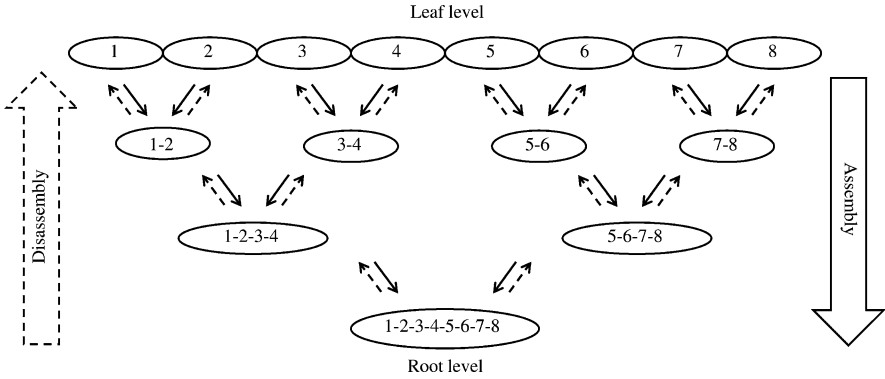
In the above equations,  $\mathcal{A}_i^k$  ( $i = 1, 2$ ) represent the spatial accelerations of the handles of the body. The terms  $\mathcal{F}_{ic}^k$  ( $i = 1, 2$ ) are the spatial constraint forces due to the kinematic constraint associated with the connecting joint, while  $\hat{f}_k$  is the column matrix of the generalized forces associated with those degrees of freedom of the joint which are represented by the normalized subspace of the joint free-motion map  $\mathcal{P}_s^{J^k}$  [57]. It is proven in [57, 60] that  $\mathcal{P}_s^{J^k} \hat{f}_k$  is the dynamically equivalent spatial force [57] due to the corresponding generalized force  $\hat{f}_k$ . As such, the handle equations of motion in a Generalized-DCA can accommodate the known/unknown generalized forces, as well as the unknown spatial forces. Different applications of these equations are reported in [57].

Similar to the DCA, the Generalized-DCA is implemented using a series of recursive assembly and disassembly processes [50] to respectively form and solve the equations of motion of the system. The main goal of the assembly pass is to recursively generate larger encompassing subsystems by assembling the adjacent articulated bodies/subsystems of a multibody system as shown in Fig. 3. It is seen that the information flow in the underlying recursive operations is not serial, but in the structure of a binary tree. In the disassembly process, these equations of motion are then solved for the spatial accelerations and constraint forces of the handles of all nodes of the binary tree of Fig. 3. The detailed information on how the new terms due to the application of the generalized forces are treated in these processes are explained in [57, 59, 60].

### 4.2 Model Adjustment

In the adaptive scheme, it may be deemed necessary to change the definition (i.e. locking or releasing) of the joints based on the value of the applied metrics. This model transition is realized by adjusting the joint free-motion map  $\mathcal{P}_s^{J^k}$  [50] of the associated joint, and the corresponding orthogonal complement map  $\mathcal{D}^{J^k}$  at the leaf level of the binary tree. For instance, if a revolute joint is to be locked because it is determined to be making an insignificant contribution to the overall conformation





**Fig. 3** Assembly and disassembly passes to recursively form and solve the equations of motion of the nodes of the binary tree

of the system, the associated spatial joint free-motion map  $\mathcal{P}^{J^k} = [1 \ 0 \ 0 \ 0 \ 0 \ 0]^T$  is replaced by the null matrix after the transition. Similarly, the current orthogonal complement of the joint free-motion map

$$D^{J^k} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

is replaced by the identity matrix if expressed in the joint coordinate system.

Using the appropriate matrices characterizing the kinematics of the joints of the desired model after the transition, the assembly and disassembly processes are then performed as described previously to form and solve the forward dynamics equations of motion of the revised system model. More detail on these model adjustments is presented in [57, 63].

### 4.3 Dynamics of the System Within the Transition

All of the adjustments between different system models including coarse-to-fine and fine-to-coarse transitions are incurred without the influence of any external load. Furthermore, the configuration of the system does not change within each model adjustment. Therefore, any violation in the conservation of the generalized momentum of the system in these transitions leads to nonphysical results. In other words, the integration of the momentum of each differential element projected onto the space of admissible motions permitted by the more restrictive model (whether pre- or post-transition) over the entire system must be conserved across the model transition [37]. For instance, consider the process of instantaneous coarse graining of

a fine-scale model. The coarse-scale model at  $t = t^c$  is realized by imposing constraints on the fine-scale model at  $t = t^f$  where  $|t^c - t^f| = \varepsilon$  where  $\varepsilon$  is vanishingly small. In this transition, the number of generalized speeds of the system reduces from  $\{u_i^f\}_{i=1}^n$  to  $\{u_i^c\}_{i=1}^{n-m}$ , where  $n$  and  $n - m$  represent the number of degrees of freedom of the fine and coarse models, respectively. As such, the conservation of the generalized momentum of the system is expressed as

$$\mathcal{L}^{c/c} = \mathcal{L}^{f/c}, \quad (5)$$

where  $\mathcal{L}$  is the generalized momentum of the system. The terms  $\mathcal{L}^{c/c}$  and  $\mathcal{L}^{f/c}$  represent the momenta of the coarse and fine models, respectively, projected onto the space of admissible motions (partial velocity vectors) of the coarse model at the time of transition.

Hence, the adaptive framework should be equipped with the machinery to efficiently form the generalized momenta balance equations and solve for the generalized speeds corresponding to the new set of degrees of freedom after each model adjustment. The formation of these impulse-momentum equations within each transition can be efficiently performed with a divide-and-conquer based scheme [52, 57]. As with the forward dynamics DCA assembly process, the handle equations which address the impulse-momentum equations of the bodies/assemblies of the system are recursively formed and combined together to find the associated equations of the resulting assemblies. These handle equations are then recursively solved for the jumps in the spatial velocities of the handles of the assemblies, as well as the constraint impulses applied to these locations when the generalized momentum balance is enforced.

## 5 Issues in Transitioning to Higher Fidelity Models

The transition towards the coarser-scale model which is effectively solving Eq. (5) for  $\{u_i^c\}_{i=1}^{n-m}$  always results in a unique solution. Unlike real mechanical systems, the conservation of the generalized momentum across the model transition to the finer-scale model of a biomolecular system may not result in a unique solution. Because, in the transitions to the coarser model, naturally existing higher modes of motion are ignored since the internal metric had previously indicated these modes as less relevant. As such, in the transition to the higher fidelity models, the kinetic energy of these ignored modes must be estimated and considered appropriately. The generalized momentum conserving distribution (Eq. (5)) of the kinetic energy among the modes of the fine model is not unique even if the value of the lost kinetic energy within this transition is known [62]. More specifically, an underdetermined set of equations must be solved for  $\{u_i^f\}_{i=1}^n$  when more than a single degree of freedom of the system is released to achieve a higher fidelity model [57].

To solve the problem of the transition to the finer fidelity model, an optimization-based technique may be used to arrive at the “best” solutions from an infinite pool of possible physically meaningful solutions. This optimization problem which is

implemented on a knowledge-, math-, or physics-based cost function is heavily constrained since it must additionally satisfy the impulse-momentum equations i.e. Eq. (5). For instance, one may minimize the  $L^2$  norm of the difference between the generalized speeds of the fine-scale model and those of the coarse-scale model. In this situation, the generalized speeds are prevented from deviating greatly from the ones before unlocking [3, 62]. Alternatively, it may be desired to perform the optimization on the energy transferred to the solvent [62].

The application of the traditional methods such as Lagrange multipliers to form and solve this constrained optimization problem is computationally expensive [57]. The computational complexity of this problem reduces via changing the constrained optimization problem to an unconstrained one which effectively reduces the optimization parameters. However, in this scheme, using the coordinate partitioning to find the relations between the dependent and independent design parameters (generalized speeds) may become very costly [57, 78] if not performed wisely. The mathematical framework of forming the impulse-momentum equations (constraints), and efficiently finding the relations between the dependent generalized speeds and the independent ones in the DCA scheme for rigid and flexible body systems are presented in [4, 41, 57, 64]. It is demonstrated in these works that the application of this algorithm can significantly reduce the computational expenses associated with the manipulations performed to derive the dynamics of the transition, as well as those performed as parts of the optimization problem.

## 6 Preliminaries for Efficient Pairwise Force Calculations

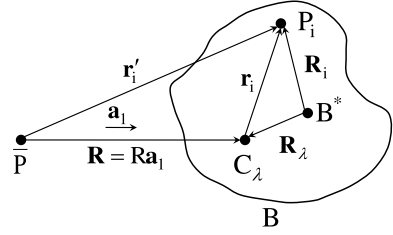
1. Consider body  $B$  (not necessarily a rigid body) containing  $N$  particles, and the individual particle  $\bar{P}$  shown in Fig. 4. In general, the pairwise force interaction between an arbitrary particle  $P_i$  embedded in  $B$  and  $\bar{P}$  may be expressed in the following format:

$$\mathbf{F}_{\bar{P}P_i} = \frac{\beta \bar{\lambda} \lambda_i}{(|\mathbf{r}'_i|)^s} \mathbf{e}_{\mathbf{r}'_i} = \beta \bar{\lambda} \lambda_i \mathbf{r}'_i (r'_i)^{-\frac{s+1}{2}}, \quad (6)$$

where,  $\beta$  is the constant associated with the force field of interest,  $s > 1$  is an integer,  $\mathbf{e}_{\mathbf{r}'_i}$  is the unit vector from  $\bar{P}$  to  $P_i$ . Additionally,  $\lambda_i$  is the quantity corresponding to the force field which is associated with particle  $P_i$ , and similarly,  $\bar{\lambda}$  is the same quantity associated with particle  $\bar{P}$ . This general formulation may be used to address the gravitational, Coulombic or London forces. For instance, if one is interested in the pairwise interactions due to the Coulomb's law [65],  $\lambda_i$  represents the charge of the particle,  $s$  becomes 2, and  $\beta$  is replaced by the Coulomb force constant.

2. For body (pseudo-atom)  $B$ , the *pseudo-center* denoted by  $C_\lambda$  is defined as the center of the body corresponding to the quantity of interest  $\lambda$  provided that the pseudo-mass (lumped quantity) of the body defined as

**Fig. 4** Pairwise interactions between the particle  $\bar{P}$  and the particles embedded in  $\bar{B}$  (particle-body interactions)



$$\Lambda \triangleq \sum_{i=1}^N \lambda_i \quad (7)$$

is not zero. The position of this point with respect to the mass center of the body (i.e.,  $B^*$ ) is calculated using the relation

$$\mathbf{R}_\lambda = \frac{\sum_{i=1}^N \mathbf{R}_i \lambda_i}{\Lambda}, \quad (8)$$

where  $\mathbf{R}_i$  is the position vector of the particle (atom)  $P_i$  measured from the center of mass of the body. Discussions on the subdomains of the system for which the pseudo-center is not defined ( $\Lambda = 0$ ) are provided in Sect. 9.

3. For body  $B$ , the *pseudo-inertia* tensor associated with the quantity  $\lambda$  with respect to the pseudo-center of the body is defined as

$$\mathcal{I}_\lambda^{B/C_\lambda} \triangleq \sum_{i=1}^N (\mathcal{U} \mathbf{r}_i^2 - \mathbf{r}_i \mathbf{r}_i) \lambda_i. \quad (9)$$

In this definition,  $\mathcal{U}$  denotes the identity tensor, and  $\mathbf{r}_i$  is the position vector of the particle  $P_i$  relative to the pseudo-center of the body. This tensor represents the dyadic of the moment of inertia (second moment) of the body if one studies the gravitational force [37]. It should be mentioned that for *rigid* subdomains of the system, this dyadic is constant if expressed in the axes fixed in the domain of interest.

## 7 Force Approximation

In the following, the approximation of the net force applied to a body due to the interactions between a single particle and the particles in the body is presented. Additionally, the resultant force interactions between the particles in two different domains (bodies) of the system is also approximated.

### 7.1 Particle-Body Force Interactions

Consider the interaction between particle  $\bar{P}$  and an arbitrary particle  $P_i$  embedded in body  $B$  shown in Fig. 4 is expressed by Eq. (6). Also assume that the origin of

the body-fixed frame is located at the pseudo-center of the body. According to the geometry shown in Fig. 4, the position vector from  $\mathbf{r}'_i$  in Eq. (6) can be replaced by  $\mathbf{R} + \mathbf{r}_i$  to arrive at

$$\mathbf{F}_{\bar{P}P_i} = \beta \bar{\lambda} \lambda_i (\mathbf{R}^2 + \mathbf{r}_i^2 + 2\mathbf{R} \cdot \mathbf{r}_i)^{-\frac{s+1}{2}} (\mathbf{R} + \mathbf{r}_i). \quad (10)$$

The above equation can be rewritten as

$$\mathbf{F}_{\bar{P}P_i} = \frac{\beta \bar{\lambda} \lambda_i}{R^s} \left( 1 + \left( \frac{r_i}{R} \right)^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right)^{-\frac{s+1}{2}} \left( \mathbf{a}_1 + \frac{\mathbf{r}_i}{R} \right), \quad (11)$$

where,  $r_i$ , and  $R$  are the lengths of the vectors  $\mathbf{r}_i$  and  $\mathbf{R}$ , respectively. In this relation,  $\mathbf{a}_1$  is the unit vector from  $\bar{P}$  to the pseudo-center of body  $B$  as shown in Fig. 4.

Using the binomial series expansion, the term  $(1 + (\frac{r_i}{R})^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R})^{-\frac{s+1}{2}}$  in Eq. (11) is expanded as

$$\begin{aligned} & \left( 1 + \left( \frac{r_i}{R} \right)^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right)^{-\frac{s+1}{2}} \\ &= 1 - \frac{s+1}{2} \left( \left( \frac{r_i}{R} \right)^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right) + \frac{(s+1)(s+3)}{8} \left( \left( \frac{r_i}{R} \right)^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right)^2 \\ & \quad - \frac{(s+1)(s+3)(s+5)}{48} \left( \left( \frac{r_i}{R} \right)^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right)^3 + \dots \end{aligned} \quad (12)$$

provided that  $|\left(\frac{r_i}{R}\right)^2 + 2\mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R}| < 1$ .

The total force experienced by body  $B$  due to the pairwise interactions between its own particles and  $\bar{P}$  is expressed as

$$\mathbf{F}_{\bar{P}B} = \sum_{i=1}^N \mathbf{F}_{\bar{P}P_i}. \quad (13)$$

Using the expression provided in Eq. (12), this net force is rewritten as

$$\begin{aligned} \mathbf{F}_{\bar{P}B} &= \frac{\beta \bar{\lambda}}{R^s} \left[ \sum_{i=1}^N \lambda_i \mathbf{a}_1 - \sum_{i=1}^N (s+1) \lambda_i \left( \mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right) \mathbf{a}_1 + \sum_{i=1}^N \lambda_i \frac{\mathbf{r}_i}{R} \right. \\ & \quad - \sum_{i=1}^N \frac{(s+1)}{2} \lambda_i \left( \frac{r_i}{R} \right)^2 \mathbf{a}_1 + \sum_{i=1}^N \frac{(s+1)(s+3)}{2} \lambda_i \left( \mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right)^2 \mathbf{a}_1 \\ & \quad \left. - \sum_{i=1}^N (s+1) \lambda_i \left( \mathbf{a}_1 \cdot \frac{\mathbf{r}_i}{R} \right) \frac{\mathbf{r}_i}{R} + O\left( \frac{r}{R} \right)^3 \right]. \end{aligned} \quad (14)$$

In the above relation,  $r$  is the length of the position vector of a generic point on  $B$  with respect to the pseudo-center of the body.

Elaborating on different terms in the above expression based on their orders with respect to  $\frac{r_i}{R}$  [57, 61], Eq. (14) provides the net force applied to the body as

$$\mathbf{F}_{\bar{P}B} = \frac{\beta\bar{\lambda}}{R^s} \left\{ \Lambda \mathbf{a}_1 + \left[ \frac{s(s+1)}{4R^2} \text{tr}(\mathcal{J}_\lambda^{B/C_\lambda}) - \frac{(s+1)(s+3)}{2R^2} \mathbf{a}_1 \cdot \mathcal{J}_\lambda^{B/C_\lambda} \cdot \mathbf{a}_1 \right] \mathbf{a}_1 + \frac{s+1}{R^2} \mathbf{a}_1 \cdot \mathcal{J}_\lambda^{B/C_\lambda} + O\left(\frac{r}{R}\right)^3 \right\}, \quad (15)$$

where  $\text{tr}(\mathcal{J}_\lambda^{B/C_\lambda})$  is the trace of the pseudo-inertia tensor. This equation can be expressed as

$$\mathbf{F}_{\bar{P}B} = \frac{\beta\bar{\lambda}\Lambda}{R^s} \left[ \mathbf{a}_1 + \sum_{i=2}^{\infty} \mathbf{f}_i \left(\frac{r}{R}\right)^i \right], \quad (16)$$

where  $\mathbf{f}_i \triangleq \mathbf{f}_i\left(\frac{r}{R}\right)^i$  is the collection of terms associated with the  $i$ th degree of  $\frac{r}{R}$ .

Ignoring the third and higher order terms in this relation, the net force is approximated as

$$\mathbf{F}_{\bar{P}B} \approx \tilde{\mathbf{F}}_{\bar{P}B} = \frac{\beta\bar{\lambda}\Lambda}{R^s} (\mathbf{a}_1 + \mathbf{f}_2), \quad (17)$$

provided that

$$\max_{i \in B} r_i \ll R. \quad (18)$$

Introducing a dextral, orthogonal set of unit vectors,  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ , and  $\mathbf{a}_3$  with the origin passing through the pseudo-center of  $B$ , and defining the elements of the pseudo-inertia tensor in  $\mathbf{a}$ -basis as

$$I_{ij} = \mathbf{a}_i \cdot \mathcal{J}_\lambda^{B/C_\lambda} \cdot \mathbf{a}_j \quad (i, j = 1, 2, 3), \quad (19)$$

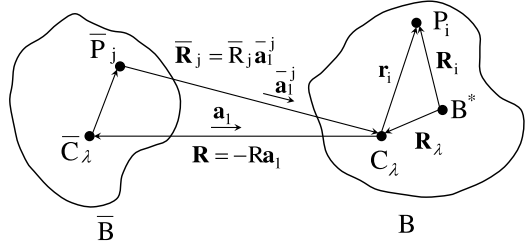
the term  $\mathbf{f}_2$  may be written as

$$\mathbf{f}_2 = \frac{s+1}{\Lambda R^2} \left\{ \left[ \frac{s}{4} \text{tr}(\mathcal{J}_\lambda^{B/C_\lambda}) - \frac{s+1}{2} I_{11} \right] \mathbf{a}_1 + I_{21} \mathbf{a}_2 + I_{31} \mathbf{a}_3 \right\}. \quad (20)$$

## 7.2 Body-Body Force Interactions

Due to the pairwise interactions between particles  $\{P_i\}_{i=1}^N$  belonging to body  $B$  (Fig. 5), and  $\{\bar{P}_j\}_{j=1}^{\bar{N}}$  embedded in body  $\bar{B}$ , the resultant force  $\mathbf{F}_{\bar{B}B}$  is applied to  $B$  by  $\bar{B}$ . This force can be approximated ( $\tilde{\mathbf{F}}_{\bar{B}B}$ ) by summing over all approximate forces applied to  $B$  by all particles  $\bar{P}_j$  on  $\bar{B}$ . As such, using the second order approximation of Eq. (15), this net force is approximated as

**Fig. 5** Pairwise interactions between the particles embedded in  $B$  and  $\bar{B}$  (body-body interaction)



$$\begin{aligned} \tilde{\mathbf{F}}_{\bar{B}B} = & \sum_{j=1}^{\bar{N}} \frac{\beta \Lambda \lambda_j}{\bar{R}_j^s} \bar{\mathbf{a}}_1^j + \frac{\beta \lambda_j}{\bar{R}_j^s} \sum_{j=1}^{\bar{N}} \left\{ \left[ \frac{s(s+1)}{4\bar{R}_j^2} \text{tr}(\mathcal{J}_\lambda^{B/C_\lambda}) \right. \right. \\ & \left. \left. - \frac{(s+1)(s+3)}{2\bar{R}_j^2} \bar{\mathbf{a}}_1^j \cdot \mathcal{J}_\lambda^{B/C_\lambda} \cdot \bar{\mathbf{a}}_1^j \right] \bar{\mathbf{a}}_1^j + \frac{s+1}{\bar{R}_j^2} \bar{\mathbf{a}}_1^j \cdot \mathcal{J}_\lambda^{B/C_\lambda} \right\}, \end{aligned} \quad (21)$$

where  $\bar{R}_j$  denotes the distance from  $\bar{P}_j$  to  $C_\lambda$ , and  $\bar{\mathbf{a}}_1^j$  is the corresponding unit vector.

Since the second summation involves the terms of second or higher degrees in  $r_i$ ,  $\bar{R}_j$  and  $\bar{\mathbf{a}}_1^j$  may be replaced by  $R$  and  $\mathbf{a}_1$ , respectively. No terms of interest for the purpose at hand are lost through this replacement. This substitution is effectively the application of the Taylor series expansion in the approximation.

Let us define  $\mathbf{a}_2$  and  $\mathbf{a}_3$  such that  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ , and  $\mathbf{a}_3$  establish a dextral, orthogonal set of unit vectors. Defining the moments and products of inertia of  $B$  and  $\bar{B}$  for axes parallel to  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ , and passing through the pseudo-center of the individual bodies as

$$I_{ij} = \mathbf{a}_i \cdot \mathcal{J}_\lambda^{B/C_\lambda} \cdot \mathbf{a}_j \quad (i, j = 1, 2, 3), \quad (22)$$

$$\bar{I}_{ij} = \mathbf{a}_i \cdot \bar{\mathcal{J}}_\lambda^{\bar{B}/\bar{C}_\lambda} \cdot \mathbf{a}_j \quad (i, j = 1, 2, 3), \quad (23)$$

and elaborating on the first summation of Eq. (21), this equation is simplified as

$$\tilde{\mathbf{F}}_{\bar{B}B} = \frac{\beta \Lambda \bar{\Lambda}}{R^s} (\mathbf{a}_1 + \mathbf{g}_2 + \bar{\mathbf{g}}_2), \quad (24)$$

where

$$\mathbf{g}_2 = \frac{s+1}{\Lambda R^2} \left\{ \left[ \frac{s}{4} \text{tr}(\mathcal{J}_\lambda^{B/C_\lambda}) - \frac{(s+1)}{2} I_{11} \right] \mathbf{a}_1 + I_{21} \mathbf{a}_2 + I_{31} \mathbf{a}_3 \right\}, \quad (25)$$

$$\bar{\mathbf{g}}_2 = \frac{s+1}{\bar{\Lambda} R^2} \left\{ \left[ \frac{s}{4} \text{tr}(\bar{\mathcal{J}}_\lambda^{\bar{B}/\bar{C}_\lambda}) - \frac{(s+1)}{2} \bar{I}_{11} \right] \mathbf{a}_1 + \bar{I}_{21} \mathbf{a}_2 + \bar{I}_{31} \mathbf{a}_3 \right\}. \quad (26)$$

## 8 Torque Approximation

Since the resultant forces calculated in Sects. 7.1 and 7.2 do not necessarily act through the center of mass of the body, they create moments about the mass center.

In the following, these resultant torques due to the long-range particle-body and body-body interactions are calculated.

### 8.1 Particle-Body Torque Interactions

Based on the geometry shown in Fig. 4, the following cross product is used to calculate the torque which body  $B$  experiences about its mass center  $B^*$  due to the interaction between  $\bar{P}$  and  $P_i$

$$\mathbf{M}_{\bar{P}P_i}^{B^*} = \mathbf{R}_i \times \mathbf{F}_{\bar{P}P_i}. \quad (27)$$

Replacing  $\mathbf{R}_i$  by  $\mathbf{R}_\lambda - \mathbf{R} + \mathbf{r}'_i$ , this torque can be rewritten as

$$\mathbf{M}_{\bar{P}P_i}^{B^*} = (\mathbf{R}_\lambda - \mathbf{R}) \times \mathbf{F}_{\bar{P}P_i} + \underbrace{\mathbf{r}'_i \times \mathbf{F}_{\bar{P}P_i}}_{\mathbf{0}}. \quad (28)$$

The last term in the above relation disappears since both  $\mathbf{r}'_i$  and  $\mathbf{F}_{\bar{P}P_i}$  are collinear vectors. As such, body  $B$  experiences the following moment about  $B^*$  due to the interactions between its own particles and  $\bar{P}$

$$\mathbf{M}_{\bar{P}B}^{B^*} = \sum_{i=1}^N (\mathbf{R}_\lambda - \mathbf{R}) \times \mathbf{F}_{\bar{P}P_i} = (\mathbf{R}_\lambda - \mathbf{R}) \times \mathbf{F}_{\bar{P}B}. \quad (29)$$

Using the second order approximation of the net force from Eq. (15), this moment is approximated as

$$\begin{aligned} \tilde{\mathbf{M}}_{\bar{P}B}^{B^*} = \frac{\beta \Lambda \bar{\lambda}}{R^s} \mathbf{R}_\lambda \times \left\{ \mathbf{a}_1 + \frac{s+1}{\bar{\lambda} R^2} \left[ \left[ \frac{s}{4} \text{tr}(\mathcal{J}_\lambda^{B/C_\lambda}) - \frac{(s+3)}{2} \mathbf{a}_1 \cdot \mathcal{J}_\lambda^{B/C_\lambda} \cdot \mathbf{a}_1 \right] \mathbf{a}_1 \right. \right. \\ \left. \left. + \mathbf{a}_1 \cdot \mathcal{J}_\lambda^{B/C_\lambda} \right] \right\} - \frac{\beta \bar{\lambda} (s+1)}{R^{(s+1)}} \mathbf{a}_1 \times (\mathbf{a}_1 \cdot \mathcal{J}_\lambda^{B/C_\lambda}). \end{aligned} \quad (30)$$

Defining the elements of the pseudo-inertia tensor in  $\mathbf{a}$ -basis as defined in Eq. (19), and using  $\mathbf{f}_2$  from Eq. (20), this expression is simplified as

$$\tilde{\mathbf{M}}_{\bar{P}B}^{B^*} = \frac{\beta \Lambda \bar{\lambda}}{R^s} \mathbf{R}_\lambda \times (\mathbf{a}_1 + \mathbf{f}_2) - \frac{\beta \bar{\lambda} (s+1)}{R^{(s+1)}} (I_{21} \mathbf{a}_3 - I_{31} \mathbf{a}_2). \quad (31)$$

### 8.2 Body-Body Torque Interactions

Using Eq. (30), the resultant moment applied to  $B$  about  $B^*$  from  $\bar{B}$  due to the interactions between the particles in these bodies can be approximated by summing over the approximate torques which body  $B$  experiences about  $B^*$  by all particles  $\bar{P}_j$  on  $\bar{B}$  as



$$\begin{aligned}
\tilde{\mathbf{M}}_{\bar{B}\bar{B}}^{B*} &= \sum_{j=1}^{\bar{N}} \frac{\beta \bar{\lambda}_j \Lambda}{\bar{R}_j^s} \mathbf{R}_\lambda \times \bar{\mathbf{a}}_1^j + \sum_{j=1}^{\bar{N}} \frac{\beta \bar{\lambda}_j}{\bar{R}_j^s} \mathbf{R}_\lambda \times \frac{s+1}{\bar{R}_j^2} \left\{ \left[ \frac{s}{4} \text{tr}(\mathcal{J}_\lambda^{B/C_\lambda}) \right. \right. \\
&\quad \left. \left. - \frac{(s+3)}{2} \bar{\mathbf{a}}_1^j \cdot \mathcal{J}_\lambda^{B/C_\lambda} \cdot \bar{\mathbf{a}}_1^j \right] \bar{\mathbf{a}}_1^j + \bar{\mathbf{a}}_1^j \cdot \mathcal{J}_\lambda^{B/C_\lambda} \right\} \\
&\quad - \sum_{j=1}^{\bar{N}} \frac{\beta \bar{\lambda}_j (s+1)}{\bar{R}_j^{(s+1)}} \bar{\mathbf{a}}_1^j \times (\bar{\mathbf{a}}_1^j \cdot \mathcal{J}_\lambda^{B/C_\lambda}). \tag{32}
\end{aligned}$$

Similar to the argument made in Sect. 7.2,  $\bar{R}_j$  and  $\bar{\mathbf{a}}_1^j$  may be replaced by  $R$  and  $\mathbf{a}_1$ , respectively. Using the same strategy provided in [61], the approximate moment about  $B^*$  from  $\bar{B}$  due to the pairwise interactions between the particles embedded in  $B$  and  $\bar{B}$  is expressed as

$$\tilde{\mathbf{M}}_{\bar{B}\bar{B}}^{B*} = \frac{\beta \Lambda \bar{\Lambda}}{R^s} \mathbf{R}_\lambda \times (\mathbf{a}_1 + \mathbf{g}_1 + \bar{\mathbf{g}}_2) - \frac{\beta \bar{\Lambda} (s+1)}{R^{(s+1)}} (I_{21} \mathbf{a}_3 - I_{31} \mathbf{a}_2), \tag{33}$$

where  $I_{ij}$ ,  $\mathbf{g}_1$ , and  $\bar{\mathbf{g}}_2$  have already been defined in Eqs. (22), (25) and (26), respectively.

## 9 Discussions on the Developed Approximations

1. The presented approximations contain the terms up to the quadrupole moment (quadrupole-quadrupole interactions). Furthermore, since the origin of the body-fixed frame is located at its pseudo-center, the first moment  $\sum_{i=1}^N \lambda_i \mathbf{r}_i$  does not appear in these approximations [57, 61].
2. According to Eq. (8), the pseudo-center is not defined when the pseudo-mass  $\Lambda$  of the body (subdomain) is zero. Similarly, if  $\Lambda$  is very close to zero, the pseudo-center may be located far away from the body. As such, the center of mass of the body may be considered as the origin of the body-fixed frame to derive the resultant forces and torques. In these situations, the pseudo-inertia tensor used in all the derived approximations is defined about  $B^*$ . Furthermore, due to locating the origin of the body-fixed frame at the mass center rather than the pseudo-center, the first moment appears in the approximations. However, it is proven in [57, 61] that the first moment is constant regardless of the choice of the origin. Consequently, when the pseudo-mass of the pseudo-atom is zero, it is necessary to express the first moment measured from any reference point defined by the analyst in the approximate force and torque. Furthermore, for rigid subdomains of the system, this term becomes time-invariant.
3. For *rigid* pseudo-atoms, the pseudo-inertia tensor which appears in all the approximations is a time-invariant quantity if expressed in body-fixed frame. As such, if this tensor is calculated for a rigid subdomain of the system at some time either before or during the simulation, no additional cost is incurred to form or use this dyadic during the course of the simulation. It is only needed to monitor the location and orientation of each pseudo-atom. Moreover, the trace of this

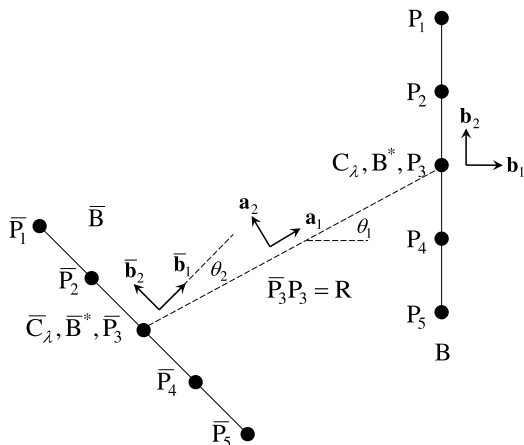
tensor which appears in these approximations is also a constant scalar for rigid subdomains.

4. Due to the symmetry observed in the body-body force approximation in Eqs. (24)–(26), the approximate net force does not violate Newton’s third law of motion.
5. In the formulation presented for the particle-body and body-body torque approximations in Eqs. (31) and (33), if the pseudo-center and the center of mass of the body coincide, i.e.,  $\mathbf{R}_\lambda = \mathbf{0}$ , the last term only contributes to the applied moment. In this case, the approximate moment formulation provides a zero value if  $\mathbf{a}_1$  is aligned with one of the principal axes of the pseudo-inertia tensor of body  $B$ . Since the geometry of the system is known at each time step, one can avoid using these approximations by checking whether the orientation of the body is close to this specific configuration. In such situations, the analyst may use the exact calculations or a higher order approximation to find the moment. Moreover, in the body-body torque interaction, one may use the summation of the particle-body torque approximations over the entire particles of the body to calculate the torque applied to a body from another body.
6. Poursina and Anderson [61] demonstrate the efficiency of the method by comparing the operation counts for particle-body and body-body force interactions using the exact force calculations and the presented approximations. Additionally, for a system containing  $n_p$  particles, and  $N_r$  rigid subdomains ( $N_r \ll n_p$ ), the computational complexity of the presented approximations is  $O(N_r)^2$  as opposed to  $O(n^2)$  complexity when all atoms are considered in pairwise calculations. Moreover, it is expected that the computational complexity will improve to  $O(N_r \log N_r)$  if advanced algorithms are used to implement these approximations [57]. As such, an efficient implementation of these approximations which is well-suited in combination with the state-of-the-art multibody algorithms is an ongoing research by the authors [57].

## 10 Numerical Results

Consider particles with unit positive charges distributed equidistantly on two straight lines  $B$  and  $\bar{B}$  with the length  $L$  as shown in Fig. 6. This could be a simple model of a rigidified residue of a DNA or an RNA. Since the Coulombic potential field is active between these charged particles, the values of the parameters in Eq. (6) are selected as  $\beta = k_e = 8.9885518 \times 10^9 \text{ N m}^2/\text{C}^2$ ,  $\lambda_i = \bar{\lambda} = \lambda = 1.6021764 \times 10^{-19} \text{ C}$ , and  $s = 2$ . Due to the symmetry in the mass and charge distribution, the mass center and the pseudo-center of each body coincide. The dextral orthogonal unit vectors  $\mathbf{b}_1\mathbf{b}_2\mathbf{b}_3$  and  $\bar{\mathbf{b}}_1\bar{\mathbf{b}}_2\bar{\mathbf{b}}_3$  are attached to the pseudo-centers of bodies  $B$  and  $\bar{B}$ , respectively. The angles  $\theta_1$  and  $\theta_2$  as shown in Fig. 6 are used to describe the orientation of these body-attached reference frames with respect to the

**Fig. 6** Pairwise Coulombic interactions between the particles on two rigid bodies  $B$  and  $\bar{B}$



**a**-frame. Further, the pseudo-inertia tensor of each body with respect to the associated pseudo-center is expressed in the corresponding body-basis as follows

$$\mathcal{J}_\lambda^{B/C_\lambda} = \mathcal{J}_\lambda^{\bar{B}/\bar{C}_\lambda} = \begin{bmatrix} \frac{5\lambda L^2}{8} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{5\lambda L^2}{8} \end{bmatrix}. \quad (34)$$

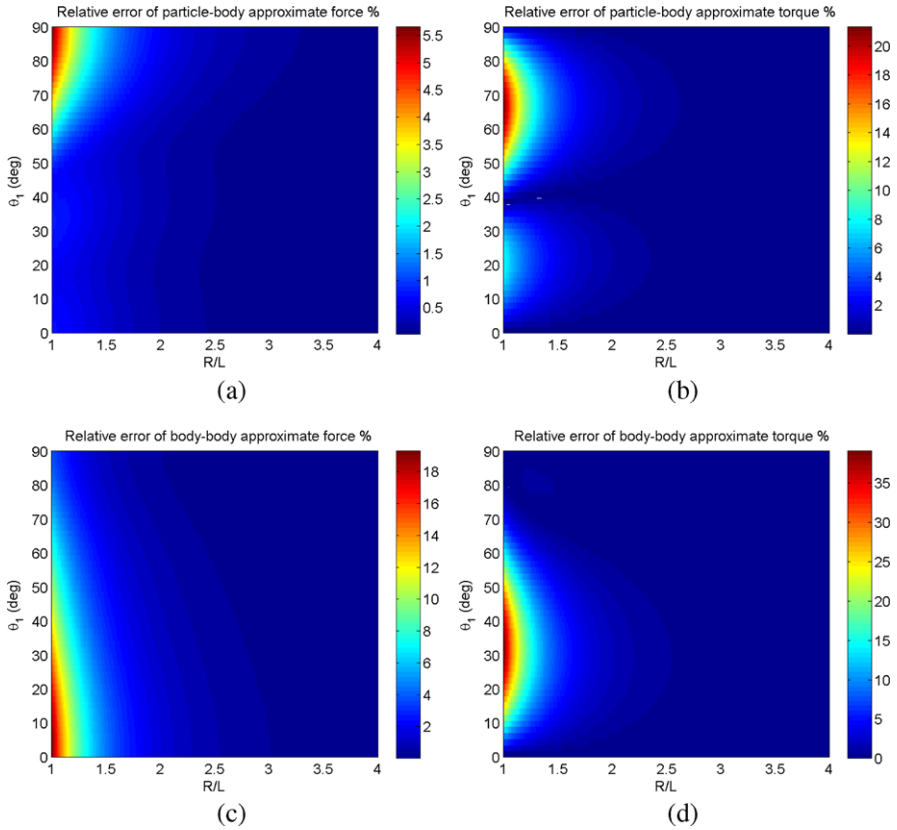
To run the simulations, it is assumed that  $L = 5 \text{ \AA}$  which is approximately on the same order of magnitude of length of an RNA residue. Herein, various configurations of this planar system are formed by fixing  $\theta_2 = 0$ , and changing  $R$  from  $L$  to  $4L$  and  $\theta_1$  from  $0$  to  $\pi/2$ . The resultant electrostatic force applied from  $\bar{B}$  to  $B$  at each configuration is computed using three different methods. First, the exact net force due to the pairwise interactions is calculated. Then, the resultant force applied by each particle of  $\bar{B}$  to the entire body  $B$  is found using the particle-body approximation derived in Eq. (17), and summed over all particles embedded in  $B$ . Finally, the resultant force applied to  $B$  is calculated using the body-body approximation presented in Eq. (24). The torque experienced by  $B$  about  $B^*$  due to the interactions between the particles embedded in  $B$  and  $\bar{B}$  is also calculated using these three methods.

Using the following definitions

$$E_{ij}^{\mathbf{F}} \triangleq \frac{\|\mathbf{F}_{approx.}(R_i, \theta_{1_j}) - \mathbf{F}_{exact}(R_i, \theta_{1_j})\|_2}{\|\mathbf{F}_{exact}(R_i, \theta_{1_j})\|_2}, \quad (35)$$

$$E_{ij}^{\mathbf{T}} \triangleq \frac{\|\mathbf{T}_{approx.}(R_i, \theta_{1_j}) - \mathbf{T}_{exact}(R_i, \theta_{1_j})\|_2}{\|\mathbf{T}_{exact}(R, \theta_1)\|_\infty}, \quad (36)$$

the values of the percentage error of the approximate resultant force and moment at each configuration  $(R_i, \theta_{1_j})$  are, respectively, calculated and depicted in Fig. 7. Since for this problem, the resultant force never becomes zero, the percentage error of the approximate force at each configuration is normalized by  $L^2$  norm of the associated force. However, the percentage error of the approximate torque is normalized

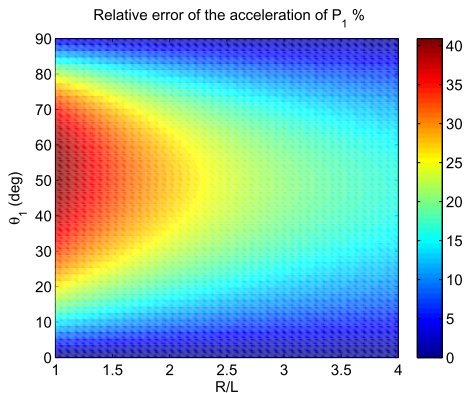


**Fig. 7** (a) The percentage error of the approximate net force applied to  $B$  by summing over particle-body approximations. (b) The percentage error of the approximate resultant torque about  $B^*$  by summing over particle-body moment approximations. (c) The percentage error of the approximate net force applied to  $B$  using body-body approximation. (d) The percentage error of the approximate resultant moment about  $B^*$  using body-body torque approximation

by the maximum of the absolute value of the exact torque among all configurations since the net moment is zero in some configurations.

The results show that both particle-body and body-body formulas generate the acceptable approximations for the far-field interactions. Although in the entire configuration space sampled in this example, there exist very tiny regions for which the approximations provide large errors, these errors decay very quickly as the bodies become more distant. For instance, for a very conservative case with  $\frac{R}{L} > 3$ , Fig. 7(c) shows that the body-body force approximation provides the relative error less than 0.1 %. It is also observed that summing over the particle-body approximation to find the force and moment between two bodies, in general, provides less error than the application of the body-body approximation. Although the particle-body approximation is more accurate, the body-body formulation is faster, and provides

**Fig. 8** The percentage relative error in the translational acceleration of  $P_1$  when the angular motion of the body is ignored



acceptable approximation for the exact resultant force and torque for most configurations.

To analyze the importance of the angular motion in the determination of the configuration of the system, assume that body  $B$  is a segment of a very long chain (of a biopolymer), and the motion is transferred to the outboard handle of this body through point  $P_1$ . Without loss of generality and for simplicity, it is assumed that the angular velocity of  $B$  (i.e.  ${}^N\boldsymbol{\omega}^B$ ) is zero. Therefore, the rotational motion of this body is reflected in its angular acceleration measured in the Newtonian frame of reference i.e.  ${}^N\boldsymbol{\alpha}^B$ . The exact translational acceleration of  $P_1$  in the Newtonian frame (i.e.  ${}^N\mathbf{a}_{exact}^{P_1}$ ) when the angular motion of the body is not ignored is expressed as

$${}^N\mathbf{a}_{exact}^{P_1} = {}^N\mathbf{a}^{B^*} + \underbrace{{}^N\boldsymbol{\omega}^B}_{\mathbf{0}} \times \left( \underbrace{{}^N\boldsymbol{\omega}^B}_{\mathbf{0}} \times \mathbf{r}^{B^*P_1} \right) + {}^N\boldsymbol{\alpha}^B \times \mathbf{r}^{B^*P_1}. \quad (37)$$

The term  ${}^N\mathbf{a}^{B^*}$  in the above relation is the translational acceleration of  $B^*$  in the Newtonian frame, and  $\mathbf{r}^{B^*P_1}$  denotes the position of  $P_1$  with respect to  $B^*$ . This relation demonstrates that both  $P_1$  and  $B^*$  have the same translational accelerations if the entire rotational motion of the body is ignored (i.e.,  ${}^N\boldsymbol{\omega}^B = {}^N\boldsymbol{\alpha}^B = 0$ ). As such, the percentage relative error of the translational acceleration of  $P_1$  when the angular motion of this segment is neglected is defined as

$$Err = \frac{\|{}^N\boldsymbol{\alpha}^B \times \mathbf{r}^{B^*P_1}\|_2}{\|{}^N\mathbf{a}_{exact}^{P_1}\|_2} \times 100. \quad (38)$$

This error is calculated and depicted in Fig. 8, assuming that the mass of each particle is 27.026 Daltons which is  $\frac{1}{5}$  of the mass of the nucleotide Adenine. It is observed that this error is significant within the majority of the configuration space. As a result, the angular motion needs to be considered in modeling biopolymers such as DNAs and RNAs in which the geometry plays an important role in determining the conformational motion of the system.

## 11 Conclusions

Herein, different strategies to improve the computational efficiency of the multi-body-based coarse-grained simulations of biopolymers have been presented. The adaptive modeling of these systems in which the model is adjusted within the course of the simulation has been developed. This scheme which is much more accurate than traditional static (time-invariant) coarse-grained models is capable of identifying the critical locations of the system to add or remove fidelity to or from the system model on-the-fly. Potential metrics to direct these model transitions have been presented. Furthermore, the important issues associated with the implementation of these instantaneous model adjustments within the course of the simulation have been addressed. Since this coarse graining strategy is realized in a multibody dynamics scheme, a Generalized Divide-and-Conquer (GDCA) which is highly modular and lends itself well to adaptivity has been presented. The method for rigid body systems is exact, non-iterative and efficient, providing a time optimal order  $\log n$  computational performance achieved with a processor optimal order  $n$  processors. Another aspect of this book chapter has been associated with the development of an efficient algorithm to approximate far-field interactions. The presented method approximates particle-body and body-body force and moment terms. The developed formulations are highly compatible with the state-of-the-art efficient multibody algorithms. The methods have provided relatively accurate results for the test case with Coulombic interactions. It has also been illustrated that the torque which is ignored in bead models plays an important role in more appropriately capturing the conformational motion of biopolymers.

**Acknowledgements** Support for this work received under National Science Foundation through award No. 0757936 is gratefully acknowledged. The author would like to thank Professor Alain Laederach from the University of North Carolina, Mr. Michael Sherman from Simbios: NIH Center for Biomedical Computation at Stanford University, and Dr. Kishor Bhalerao from the University of Melbourne for several useful discussions. They also would like to thank Jeremy Laffin who independently validated some of the simulation results.

## References

1. Abagyan, R., Mazur, A.: New methodology for computer-aided modeling of biomolecular structure and dynamics. 2. Local deformations and cycles. *J. Comput. Phys.* **6**(4), 833–845 (1989)
2. Anderson, K.S.: Recursive derivation of explicit equations of motion for efficient dynamic/control simulation of large multibody systems. Ph.D. thesis, Stanford University (1990)
3. Anderson, K.S., Poursina, M.: Energy concern in biomolecular simulations with transition from a coarse to a fine model. In: Proceedings of the Seventh International Conference on Multibody Systems, Nonlinear Dynamics and Control, ASME Design Engineering Technical Conference 2009 (IDETC09), IDETC2009/MSND-87297, San Diego, CA (2009)
4. Anderson, K.S., Poursina, M.: Optimization problem in biomolecular simulations with DCA-based modeling of transition from a coarse to a fine fidelity. In: Proceedings of the Seventh International Conference on Multibody Systems, Nonlinear Dynamics and Control, ASME Design Engineering Technical Conference 2009 (IDETC09), IDETC2009/MSND-87319, San Diego, CA (2009)

5. Appel, A.W.: An efficient program for many-body simulation. *SIAM J. Sci. Stat. Comput.* **6**(1), 85–103 (1985)
6. Armstrong, W.W.: Recursive solution to the equations of motion of an  $n$ -link manipulator. In: Fifth World Congress on the Theory of Machines and Mechanisms, vol. 2, pp. 1342–1346 (1979)
7. Avello, A., Jiménez, J.M., Bayo, E., García de Jalón, J.: A simple and highly parallelizable method for real-time dynamic simulation based on velocity transformations. *Comput. Methods Appl. Mech. Eng.* **107**(3), 313–339 (1993)
8. Bae, D.S., Haug, E.J.: A recursive formation for constrained mechanical system dynamics: Part I. Open loop systems. *Mech. Struct. Mach.* **15**(3), 359–382 (1987)
9. Bae, D.S., Kuhl, J.G., Haug, E.J.: A recursive formulation for constrained mechanical system dynamics: Part III. Parallel processor implementation. *Mech. Based Des. Struct. Mach.* **16**(2), 249–269 (1988)
10. Barns, J., Hut, P.: A hierarchical  $o(n \log n)$  force-calculation algorithm. *Lett. Nature* **324**(4), 446–449 (1986)
11. Becker, N.B., Everaers, R.: From rigid base pairs to semiflexible polymers: coarse-graining DNA. *Phys. Rev. E* **76**(2), 021923 (2007)
12. Bhalerao, K., Anderson, K.: Modeling intermittent contact for flexible multibody-rigid-body dynamics. *Nonlinear Dyn.* **60**(1–2), 63–79 (2010)
13. Bhalerao, K.D.: On methods for efficient and accurate design and simulation of multibody systems. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy (2010)
14. Bhalerao, K.D., Poursina, M., Anderson, K.S.: An efficient direct differentiation approach for sensitivity analysis of flexible multibody systems. *Multibody Syst. Dyn.* **23**(2), 121–140 (2010)
15. Brandl, H., Johanni, R., Otter, M.: A very efficient algorithm for the simulation of robots and similar multibody systems without inversion of the mass matrix. In: IFAC/IFIP/IMACS Symposium, Vienna, Austria, pp. 95–100 (1986)
16. Chakrabarty, A., Cagin, T.: Coarse grain modeling of polyimide copolymers. *Polymer* **51**(12), 2786–2794 (2010)
17. Chen, S.J.: RNA folding: conformational statistics, folding kinetics, and ion electrostatics. *Annu. Rev. Biophys.* **37**(1), 197–214 (2008)
18. Chirikjian, G.S., Wang, Y.: Conformational statistics of stiff macromolecules as solutions to partial differential equations on the rotation and motion groups. *Phys. Rev. E* **62**(1), 880–892 (2000)
19. Chun, H.M., Padilla, C.E., Chin, D.N., Watanabe, M., Karlov, V.I., Alper, H.E., Soosaar, K., Blair, K.B., Becker, O.M., Caves, L.S.D., Nagle, R., Haney, D.N., Farmer, B.L.: MBO(N)D: a multibody method for long-time molecular dynamics simulations. *J. Comput. Chem.* **21**(3), 159–184 (2000)
20. Chung, S., Haug, E.J.: Real-time simulation of multibody dynamics on shared memory multi-processors. *J. Dyn. Syst. Meas. Control* **115**(4), 627–637 (1993)
21. de Leeuw, S., Perram, J., Smith, E.: Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc. R. Soc. Lond. Ser. A* **373**(1752), 27–56 (1980)
22. Dill, K.A., Ozkan, S.B., Shell, M.S., Weikl, T.R.: The protein folding problem. *Annu. Rev. Biophys.* **37**(1), 289–316 (2008)
23. Ding, H.Q., Karasawa, N., Goddard, W.A. III: Atomic level simulation on a million particles: the cell multipole method for Coulomb and London nonbond interactions. *J. Chem. Phys.* **97**(6), 4309–4315 (1992)
24. Ewald, P.: Evaluation of optical and electrostatic lattice potentials. *Ann. Phys.* **64**, 253–287 (1921)
25. Featherstone, R.: The calculation of robotic dynamics using articulated body inertias. *Int. J. Robot. Res.* **2**(1), 13–30 (1983)
26. Featherstone, R.: *Robot Dynamics Algorithms*. Kluwer Academic, Boston (1987)

27. Featherstone, R.: A divide-and-conquer articulated body algorithm for parallel  $O(\log(n))$  calculation of rigid body dynamics. Part 1: Basic algorithm. *Int. J. Robot. Res.* **18**(9), 867–875 (1999)
28. Featherstone, R.: A divide-and-conquer articulated body algorithm for parallel  $O(\log(n))$  calculation of rigid body dynamics. Part 2: Trees, loops, and accuracy. *Int. J. Robot. Res.* **18**(9), 876–892 (1999)
29. Fijany, A., Bejczy, A.K.: Techniques for parallel computation of mechanical manipulator dynamics. Part II: Forward dynamics. In: Leondes, C. (ed.) *Advances in Robotic Systems and Control*, vol. 40, pp. 357–410. Academic Press, San Diego (1991)
30. Fijany, A., Sharf, I., D’Eleuterio, G.M.T.: Parallel  $O(\log n)$  algorithms for computation of manipulator forward dynamics. *IEEE Trans. Robot. Autom.* **11**(3), 389–400 (1995)
31. Freddolino, P.L., Arkhipov, A., Shih, A.Y., Yin, Y., Chen, Z., Schulten, K.: Application of residue-based and shape-based coarse graining to biomolecular simulations. In: Voith, G.A. (ed.) *Coarse-Graining of Condensed Phase and Biomolecular Systems*, pp. 299–315. CRC Press, Boca Raton (2008)
32. Greengard, L., Rokhlin, V.: A fast algorithm for particle simulations. *J. Comput. Phys.* **135**(2), 280–292 (1997)
33. Haile, J.: *Molecular Dynamics Simulation: Elementary Methods*. Wiley Interscience, New York (1992)
34. Hwang, R.S., Bae, D., Kuhl, J.G., Haug, E.J.: Parallel processing for real-time dynamics systems simulations. *J. Mech. Des.* **112**(4), 520–528 (1990)
35. Jain, A.: Unified formulation of dynamics for serial rigid multibody systems. *J. Guid. Control Dyn.* **14**(3), 531–542 (1991)
36. Jain, A., Vaidehi, N., Rodriguez, G.: A fast recursive algorithm for molecular dynamics simulation. *J. Comput. Phys.* **106**(2), 258–268 (1993)
37. Kane, T.R., Levinson, D.A.: *Dynamics: Theory and Application*. McGraw-Hill, New York (1985)
38. Kane, T.R., Likins, P.W., Levinson, D.A.: *Spacecraft Dynamics*. McGraw-Hill, New York (1983)
39. Kasahara, H., Fujii, H., Iwata, M.: Parallel processing of robot motion simulation. In: *Proceedings IFAC 10th World Conference* (1987)
40. Kazerounian, K., Latif, K., Alvarado, C.: Protofold: a successive kinetostatic compliance method for protein conformation prediction. *J. Mech. Des.* **127**(4), 712–718 (2005)
41. Khan, I., Poursina, M., Anderson, K.S.: DCA-based optimization in transitioning to finer models in articulated multi-flexible-body modeling of biopolymers. In: *Proceedings of the ECCOMAS Thematic Conference—Multibody Systems Dynamics*, Brussels, Belgium (2011)
42. Lathrop, L.H.: Parallelism in manipulator dynamics. *Int. J. Robot. Res.* **4**(2), 80–102 (1985)
43. Leach, A.R.: *Molecular Modelling Principles and Applications*, 2nd edn. Prentice Hall, New York (2001)
44. Lebrun, A., Lavery, R.: Modeling the mechanics of a DNA oligomer. *J. Biomol. Struct. Dyn.* **16**(3), 593–604 (1998)
45. Luh, J.S.Y., Walker, M.W., Paul, R.P.C.: On-line computational scheme for mechanical manipulators. *J. Dyn. Syst. Meas. Control* **102**(2), 69–76 (1980)
46. Malczyk, P., Fraczek, J.: Lagrange multipliers based divide and conquer algorithm for dynamics of general multibody systems. In: *Proceedings of the ECCOMAS Thematic Conference—Multibody Systems Dynamics*, Warsaw, Poland (2009)
47. Malczyk, P., Fraczek, J.C.: Parallel index-3 formulation for real-time multibody dynamics simulations. In: *Proceedings of the 1st Joint International Conference on Multibody System Dynamics*, Lappeenranta, Finland (2010)
48. Marrink, S.J., de Vries, A.H., Mark, A.E.: Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **108**, 750–760 (2004)
49. Mukherjee, R., Anderson, K.S.: A logarithmic complexity divide-and-conquer algorithm for multi-flexible articulated body systems. *J. Comput. Nonlinear Dyn.* **2**(1), 10–21 (2007)



50. Mukherjee, R., Anderson, K.S.: An orthogonal complement based divide-and-conquer algorithm for constrained multibody systems. *Nonlinear Dyn.* **48**(1–2), 199–215 (2007)
51. Mukherjee, R.M.: Multibody dynamics algorithm development and multiscale modelling of materials and biomolecular systems. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy (2007)
52. Mukherjee, R.M., Anderson, K.S.: Efficient methodology for multibody simulations with discontinuous changes in system definition. *Multibody Syst. Dyn.* **18**(2), 145–168 (2007)
53. Mukherjee, R.M., Bhalerao, K.D., Anderson, K.S.: A divide-and-conquer direct differentiation approach for multibody system sensitivity analysis. *Struct. Multidiscip. Optim.* **35**(5), 413–429 (2007)
54. Neilan, P.E.: Efficient computer simulation of motions of multibody systems. Ph.D. thesis, Stanford University (1986)
55. Niedermeier, C., Tavan, P.: A structure adapted multipole method for electrostatic interactions in protein dynamics. *J. Chem. Phys.* **101**(1), 734–748 (1994)
56. Parisien, M., Major, F.: The MC-fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**(7183), 51–55 (2008)
57. Poursina, M.: Robust framework for the adaptive multiscale modeling of biopolymers. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy (2011)
58. Poursina, M., Anderson, K.S.: Constant temperature simulation of articulated polymers using divide-and-conquer algorithm. In: *Proceedings of the ECCOMAS Thematic Conference—Multibody Systems Dynamics*, Brussels, Belgium (2011)
59. Poursina, M., Anderson, K.S.: Multibody dynamics in generalized divide and conquer algorithm (GDCA) scheme. In: *Proceedings of the Eighths International Conference on Multibody Systems, Nonlinear Dynamics and Control*, ASME Design Engineering Technical Conference 2011 (IDETC11), DETC2011-48383, Washington, DC (2011)
60. Poursina, M., Anderson, K.S.: An extended divide-and-conquer algorithm for a generalized class of multibody constraints. *Multibody Syst. Dyn.* (2012). doi:[10.1007/s11044-012-9324-9](https://doi.org/10.1007/s11044-012-9324-9)
61. Poursina, M., Anderson, K.S.: Long-range force and moment in multiresolution simulation of molecular systems. *J. Comput. Phys.* **231**(21), 7237–7254 (2012)
62. Poursina, M., Bhalerao, K.D., Anderson, K.S.: Energy concern in biomolecular simulations with discontinuous changes in system definition. In: *Proceedings of the ECCOMAS Thematic Conference—Multibody Systems Dynamics*, Warsaw, Poland (2009)
63. Poursina, M., Bhalerao, K.D., Flores, S., Anderson, K.S., Laederach, A.: Strategies for articulated multibody-based adaptive coarse grain simulation of RNA. *Methods Enzymol.* **487**, 73–98 (2011)
64. Poursina, M., Khan, I., Anderson, K.S.: Model transitions and optimization problem in multi-flexible-body modeling of biopolymers. In: *Proceedings of the Eighths International Conference on Multibody Systems, Nonlinear Dynamics and Control*, ASME Design Engineering Technical Conference 2011 (IDETC11), DETC2011-48383, Washington, DC (2011)
65. Poursina, M., Laffin, J., Anderson, K.S.: Fast electrostatic force and moment calculations in multibody-based simulations of coarse-grained biopolymers. In: *Proceedings of the Eighths International Conference on Multibody Systems, Nonlinear Dynamics and Control*, ASME Design Engineering Technical Conference 2011 (IDETC11), DETC2011-48383, Washington, DC (2011)
66. Praprotnik, M., Site, L., Kremer, K.: Adaptive resolution molecular-dynamics simulation: changing the degrees of freedom on the fly. *J. Chem. Phys.* **123**(22), 224106–224114 (2005)
67. Redon, S., Lin, M.C.: An efficient, error-bounded approximation algorithm for simulating quasi-statics of complex linkages. *Comput. Aided Des.* **38**(4), 300–314 (2006)
68. Rosenthal, D.: An order  $n$  formulation for robotic systems. *J. Astronaut. Sci.* **38**(4), 511–529 (1990)
69. Rosenthal, D.E., Sherman, M.A.: High performance multibody simulations via symbolic equation manipulation and Kane’s method. *J. Astronaut. Sci.* **34**(3), 223–239 (1986)
70. Rossi, R., Isorce, M., Morin, S., Flocard, J., Arumugam, K., Crouzy, S., Vivaudou, M., Redon, S.: Adaptive torsion-angle quasi-statics: a general simulation method with applications to pro-

- tein structure analysis and design. In: ISMB/ECCB (Supplement of Bioinformatics), vol. 23, pp. 408–417 (2007)
71. Scheraga, H.A., Khalili, M., Liwo, A.: Protein-folding dynamics: overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* **58**(1), 57–83 (2007)
  72. Shahbazi, Z., Ilies, H., Kazerounian, K.: Hydrogen bonds and kinematic mobility of protein molecules. *J. Mech. Robot.* **2**(2), 021009 (2010)
  73. Shillcocka, J.C., Lipowsky, R.: Equilibrium structure and lateral stress distribution of amphiphilic bilayers from dissipative particle dynamics simulations. *J. Chem. Phys.* **117**, 5048–5061 (2002)
  74. Tozzini, V.: Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* **15**(2), 144–150 (2005)
  75. Vaidehi, N., Jain, A., Goddard, W.A. III: Constant temperature constrained molecular dynamics: the Newton-Euler inverse mass operator method. *J. Phys. Chem.* **100**(25), 10508–10517 (1996)
  76. Vereshchagin, A.F.: Computer simulation of the dynamics of complicated mechanisms of robot-manipulators. *Eng. Cybern.* **12**(6), 65–70 (1974)
  77. Walker, M.W., Orin, D.E.: Efficient dynamic computer simulation of robotic mechanisms. *J. Dyn. Syst. Meas. Control* **104**(3), 205–211 (1982)
  78. Wehage, R.A., Haug, E.: Generalized coordinate partitioning for dimension reduction in analysis of constrained dynamic systems. *J. Mech. Des.* **104**, 247–255 (1982)