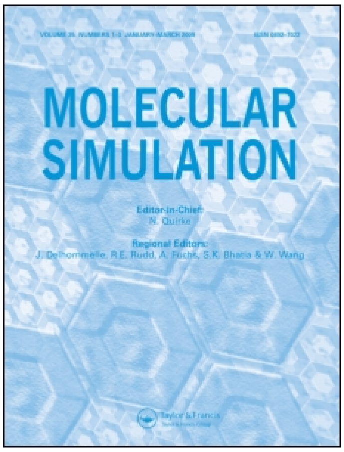

Molecular Simulation

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Solving the equations of motion for mixed atomistic and coarse-grained systems

Jong Hyuk Park[a]; Andreas Heyden[a]
[a] Department of Chemical Engineering, University of South Carolina, Columbia, SC, USA

To cite this Article Park, Jong Hyuk and Heyden, Andreas(2009) 'Solving the equations of motion for mixed atomistic and coarse-grained systems', Molecular Simulation, 35: 10, 962 — 973

To link to this Article: DOI: 10.1080/08927020902902759

URL: http://dx.doi.org/10.1080/08927020902902759

Molecular Simulation
Vol. 35, Nos. 10–11, September 2009, 962–973

Solving the equations of motion for mixed atomistic and coarse-grained systems

Jong Hyuk Park and Andreas Heyden*

Department of Chemical Engineering, University of South Carolina, 301 South Main Street, Columbia, SC 29208, USA

(*Received 8 January 2009; final version received 6 March 2009*)

A mixed-resolution molecular dynamics technique is presented that permits simulation of large and complex systems that critically depend on a subtle interplay between energy and entropy and, therefore, require both an accurate energy evaluation and an extensive sampling of a large and complex phase space. The computational approach is based upon the idea that many complex systems can be spatially divided into a relatively small active zone (that requires an accurate energy evaluation but, due to its relatively small size, only a limited sampling of phase space) and a relatively large inactive zone (that requires a less accurate energy description but the sampling of a very large phase space). Mixed-resolution models that use an accurate atomistic model for the active zone and an efficient coarse-grained model (that lumps groups of atoms into pseudo-atoms) for the inactive zone are ideally suited for these systems. One challenge in mixed-resolution simulations is creating a seamless connection between low- and high-resolution zones that exchange groups of atoms. We derive a mixed-resolution Hamiltonian and present a microcanonical simulation protocol that conserves energy and momentum and allows for a change in resolution of selected groups of atoms during a simulation. The method is applicable to simulations in other ensembles and for systems with multiple high-resolution zones. To illustrate the numerical stability of our technique, we present simulation results for a system of mixed-resolution methane molecules.

Keywords: multiscale; simulation; coarse-graining; symplectic; dynamics

1. Introduction

All-atom (AA) molecular dynamics (MD) simulations [1,2] of condensed phases and polymeric and biomolecular systems have resulted in significant insights into the molecular-scale phenomena that control these complex systems. However, many processes that occur in liquids, soft materials and biomolecular systems happen over time and length scales that are well beyond the current capabilities of AA simulations even for the most powerful parallel soft- and hardware. As a result, new computational approaches that permit the statistically meaningful investigation of longer time and length scale phenomena continue to be developed. One currently popular approach is the coarse-grained (CG) simulation which is based upon the idea that atomistic resolution is often not required for understanding phenomena that occur over millisecond time scales and hundreds of nanometre length scales. CG models lump a group of atoms into a pseudo-atom whose motion is governed by a simplified potential. Thus, they require the evaluation of fewer interactions [3] that are often shorter ranged and ‘softer’ than atomic interactions [4,5]. Since this permits the use of longer and more efficient time steps in a MD simulation [6,7], CG simulations are overall often multiple orders of magnitude more efficient than AA simulations. The main challenges that emerge in CG simulations are (i) the determination of an accurate CG potential energy function [8] and (ii) the intrinsic limited transferability of CG models. As a result, the predictive power of CG methods has been questioned [9]. To understand the challenges in obtaining reliable and predictive CG models, it is important to note that many processes depend on free energy differences so that a CG potential energy function should satisfy

$$\begin{aligned}\exp(-F/k_{\mathrm{B}}T) &= (\mathrm{const.})\int \mathrm{d}\mathbf{q}_{\mathrm{AA}}\exp(-V_{\mathrm{AA}}(\mathbf{q}_{\mathrm{AA}})/k_{\mathrm{B}}T)\\ &\approx (\mathrm{const.}')\int \mathrm{d}\mathbf{q}_{\mathrm{CG}}\exp(-V_{\mathrm{CG}}(\mathbf{q}_{\mathrm{CG}})/k_{\mathrm{B}}T), \qquad (1)\end{aligned}$$

where F is the Helmholtz free energy of the system, $V_{\mathrm{AA}}(\mathbf{q}_{\mathrm{AA}})$ is the atomistic system potential energy, $V_{\mathrm{CG}}(\mathbf{q}_{\mathrm{CG}})$ is the CG system potential energy, T is the thermodynamic temperature, k_{B} is the Boltzmann constant, and (const.) and (const.$'$) are normalisation constants. Equation (1) illustrates that certain degrees of freedom have been integrated over during the transformation from AA coordinates, $\mathbf{q}_{\mathrm{AA}}$, to CG coordinates, $\mathbf{q}_{\mathrm{CG}}$, and that therefore, the CG potential is actually a free energy (the so-called many-body potential of mean force) in CG variables that includes some entropic contributions. Consequently, CG potential energies are in principle,

---

*Corresponding author. Email: heyden@cec.sc.edu

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927020902902759
http://www.informaworld.com

thermodynamic-state dependent and should be functions of the system size (number of atoms), system volume and system temperature. The functional dependence of the CG potential energy on these state variables is currently unknown and, furthermore, system dependent. As a result, Izvekov and Voth [5,10] and Noid et al. [11,12] recently presented a multiscale approach that determines the CG potential from an atomistic simulation. While very powerful, such a multiscale approach that requires an atomistic simulation that samples the relevant phase space, is computationally expensive and has, unfortunately, often only limited predictive power.

To combine the efficiency of CG models with the accuracy of atomistic models for systems that require atomistic resolution only locally, for example at a reactive group, defect or interface (called the active zone), mixed-resolution models have been developed. These models use a low-resolution description for the part of the system distant from an active site and an atomistic description for the active site and its direct environment [13–33]. The low-resolution description can be a CG model or a mixture of CG and atomistic models.

We stress that these mixed-resolution models have the potential to be significantly more predictive than CG models while being computationally nearly as efficient as CG models. This advantage originates from the fact that CG groups in the mixed-resolution systems are usually only indirectly interacting with the active site or interface and are directly surrounded only by a relatively homogeneous medium (e.g. CG water molecules in protein-water simulations are only surrounded by CG and AA water molecules or CG polymer units in a polymer–solid surface system are only surrounded by CG and AA polymer units). As a result, the accuracy in the description of the interaction between the active site and CG group is less critical and it seems reasonable to assume that a mixed-resolution force field obtained from, e.g. water or polymer simulations, can also be used (perhaps with minor modifications) in these more complex systems that possess interfaces and active sites.

Since the active zone may diffuse during a molecular simulation or ligands may exchange in and out of the inner coordination shell of the active site, the mixed-resolution molecular simulation technique needs to permit an on-the-fly reclassification (from atomistic to low-resolution or vice versa) of atoms or groups as they transition between the high- and low-resolution regimes.

Previously, only the AdResS algorithm from Praprotnik et al. [17,21–23] and the hybrid scheme from Ensing et al. [26] permitted such a reclassification during an MD simulation. The AdResS scheme defines an atomistic zone, a CG zone and a mixed-resolution zone (in which the degrees of freedom are slowly switched on or off) in the simulation box and considers the reclassification of atoms or groups as a 'geometry-induced first-order phase transition'. Using different force fields in each zone is equivalent to using a non-conservative force field for the whole system.[1] As a result, there exists no potential energy function for the entire system, and simulations do not conserve either energy or angular momentum. In fact, in order to obtain reasonable simulation results, special thermostats have to be designed that add or remove significant amounts of energy to the groups of atoms that change their resolution. The hybrid scheme from Ensing et al. integrates, in essence, the same force field as the AdResS scheme (a different smoothing function is used), and adds or subtracts energy terms to the system energy to approximately conserve energy whenever a group crosses a zone boundary. Considering that adding or subtracting energy terms to the system energy does not alter the trajectory (the trajectory is obtained by integrating Newton's second law, $F = ma$), the hybrid scheme from Ensing et al. does not, in principle, remove any of the limitations of the AdResS algorithm. Overall, the AdResS and hybrid schemes seem to be very practical methods that are relatively easy to implement in any MD code but that can formally only describe non-conservative systems with non-conservative forces.

In this paper, we review our novel approach [34] for simulating mixed-resolution systems and solving the equations of motion for a conservative mixed atomistic and CG system. The method is called adaptive partitioning (AP) of the Lagrangian (APL). It is more elaborate to implement in any MD code than previous methods, but it is fully based on classical and statistical mechanics of conservative systems. As a result, a microcanonical-ensemble (NVE) simulation protocol conserves energy, angular and linear momentum; Newton's third law is fulfilled; and the method is also well defined for any other thermodynamic ensemble. After having described the theory of the APL method, we present an efficient explicit symplectic integrator for mixed-resolution systems, discuss the main challenges in mixed-resolution simulations, and conclude by presenting the first mixed-resolution simulation results that prove the numerical applicability of our technique.

2. Theory

This section is organised as follows: we first explain why the equations of motion of a conservative mixed-resolution system are not given by Newton's second law but by the more general Euler–Lagrange equations. Then, we describe our procedure for obtaining a smooth mixed-resolution Lagrangian and Hamiltonian function. Finally, we conclude by illustrating how to extend our method to multiple active zones and various thermodynamic ensembles.

2.1 Why the equations of motion are not given by Newton's second law

MD simulations of AA and CG systems usually involve integrating Newton's second law,

$$m_i \ddot{q}_i = F_i = -\frac{\partial V}{\partial q_i}, \tag{2}$$

where m_i is the mass of atom or CG site i, q_i is a generalised (often a Cartesian) coordinate of site i, $\ddot{q}_i$ is the second time derivative of q_i, V is the system potential energy, and F_i is the force acting on site i. Equation (2) can be derived from the more general Euler–Lagrange equations that describe the evolution of any classical system, if we assume that the Lagrangian, $L = T - V$, is not an explicit function of time, the system potential energy, V, is only a function of the generalised coordinates, q_i, and the system kinetic energy, T, is only a function of the first time derivative of the generalised coordinates, $\dot{q}_i$ (i.e. the generalised velocities) [35]. These assumptions are usually fulfilled for most molecular simulation systems, where the system kinetic energy is usually given by $T = \sum_i (1/2) m_i \dot{q}_i^2$.

In contrast, in mixed-resolution systems, the resolution of a group of atoms changes depending on the coordinates of the group of atoms (high resolution close to an active zone and low resolution far away from an active zone). Consequently, the degrees of freedom of a group of atoms also change depending on the coordinates of the group, and the kinetic energy functional of the system has to be both velocity and coordinate dependent. It is well known that for coordinate-dependent kinetic energies, Newton's second law cannot be derived from the Euler–Lagrange equations and is not valid for conservative systems [35]. The equations of motion of a conservative mixed-resolution system that allows for a change in resolution of groups of atoms and in which resolution changes are specified by coordinates are given by the Euler–Lagrange equations or the equivalent Hamilton equations. It is the main task of this section and the key idea of the APL method to design a single physically meaningful and smooth Lagrangian function that can be used to derive a conservative Hamiltonian for a mixed-resolution system.

2.2 Adaptive partitioning of Lagrangian functions

The APL method starts with an atomistic description of a system and classifies all atoms into groups with group coordinates that become the interaction sites for those groups that are CG in the low-resolution description of the system. In the following, we limit our discussion to group coordinates (CG sites) that are the centres of mass of the atoms the specific group is composed of. This choice is motivated by Noid et al.'s observation that a sufficient condition for the development of a 'consistent' CG model is that the CG site is positioned at the centre of mass of their underlying atomistic structure (assuming the CG site has a CG mass equal to the sum of their underlying atomistic masses) [11,12]. 'Consistent' in this context means that the low-resolution (CG) model samples the momentum space with the same probability as the centre of mass of the high-resolution (AA) model.

Next, we define a particular atom or a group of atoms, called an active site, as the centre of an active zone that needs to be modelled atomistically. Although the APL scheme can involve multiple active zones and each active zone can be defined very generally, we limit our discussion here to a single spherical active zone surrounded by a spherical-shell buffer zone (defined with inner and outer radii, $r_{\min}$ and $r_{\max}$) and an environmental zone (which comprises the rest of the system, which need not be spherical). Extensions to systems with multiple active zones are described in Section 2.4. At any time, all of the atoms of a given group are always considered to be in the same zone specified by the location of the group coordinates. Groups in the active zone and the buffer zone possess an explicit atomistic structure. In contrast, groups of atoms located in the environmental zone are described only by their low-resolution interaction sites. The atomistic structure of these groups does not have to be specified, and only the group coordinates have to be stored. The possibility of using an all-atomistic resolution in all zones (the force-field has different resolutions) is discussed in Section 3.

The purpose of the buffer zone is to be able to design a smooth Lagrangian function or for that matter, a smooth kinetic and potential energy of the system when groups are entering or leaving the active zone. A smooth kinetic and potential energy are necessary requirements for obtaining stable long time trajectories [1]. The thickness of the buffer depends on the numerical integration time step and the low-resolution and atomistic description but might be, for example, a few angstroms. Figure 1 illustrates the partitioning of a system into three zones for a lithium ion–water system with the lithium ion specifying the active zone in which water molecules have atomistic resolution, the buffer zone in which water molecules have a partially atomistic and partially CG resolution, and the environmental zone in which the water molecules have a CG resolution. For illustrative purposes, a single water molecule has been CG into a single CG site.

Next, we introduce the first key assumption of the APL method. We assume that it is possible to design a physically meaningful mixed-resolution Lagrangian for a fixed number of v CG and μ atomistic groups. To keep the notation simple, we will write in the following the equations for the case where every group has the same number A of atoms, but the generalisation to A_α atoms in group α is straightforward, i.e.

$$L^{(\mu,v)} = \sum_{\alpha=1}^{v} \frac{1}{2} M_\alpha \mathbf{U}_\alpha^2 + \sum_{k=1}^{A\mu} \frac{1}{2} m_k \mathbf{u}_k^2 - V^{(\mu,v)}, \tag{3}$$

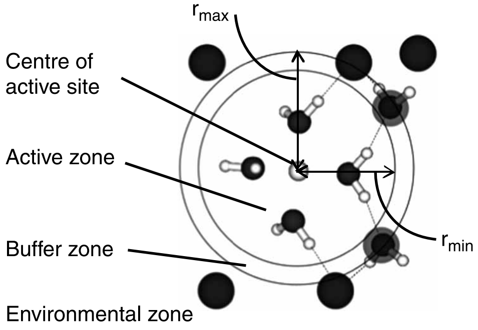

Figure 1. Partitioning of a Li^+-ion/water system into an active zone specified by the Li^+-ion, an environmental zone, and a buffer zone separating the active from the environmental zone. Groups of atoms, such as water molecules, in the active zone have a high (atomistic) resolution, groups in the environmental zone have a low (CG) resolution, and groups in the buffer zone have a mixed resolution.

where $M_\alpha = \sum_{k=1}^{A} m_k^{(\alpha)}$ is the mass of group α, $\mathbf{U}_\alpha$ is the velocity of its group coordinate, $m_k^{(\alpha)}$ is the mass of atom k in group α, $\mathbf{u}_k$ is its velocity vector $(u_{kx}, u_{ky}, u_{kz})^{\mathrm{T}}$, and $V^{(\mu,v)}$ is the potential energy describing all interactions between the low-resolution and atomistic sites. The assumption of being able to design a physically meaningful Lagrangian, $L^{(\mu,v)}$, depends heavily on being able to determine a meaningful mixed-resolution potential energy, $V^{(\mu,v)}$. Previously published simulation results from various research groups suggest that this is possible [13–33]; but we will further discuss this issue in Section 4.

The second assumption of the APL method is that if, at one instant in time during a molecular simulation, there are M groups of atoms in the active zone, K groups of atoms in the buffer zone, and N groups of atoms in the environmental zone, then we can define a physically reasonable mixed-resolution Lagrangian, L^{AP}, as a linear combination of all possible combinations of Lagrangian functions that are obtained by treating the active zone and a subset of the K groups in the buffer zone at a high level of resolution (and the rest at a low level of resolution),

$$L^{\mathrm{AP}} = L^{\mathrm{AP}}(L^{[0]}, \{L^{[1]}\}, \ldots, L^{[K]}), \tag{4}$$

where $\{L^{[J]}\}$ denotes all $K!/J!(K-J)!$ possible Lagrangian functions that can be obtained by considering various choices of J buffer-zone groups to be atomistic. Specifically, we assume that a meaningful Lagrangian function, L^{AP}, can be obtained by extending our previously developed AP method for multilevel (e.g. QM/MM) potential energies [36] to high- and low-resolution Lagrangian functions. The AP procedure guarantees that a smooth Lagrangian function, L^{AP}, is obtained when a group of atoms enters or leaves the buffer zone and changes its resolution [34,36].

To further simplify the notation, we will denote the Lagrangian as $L_{\alpha\beta\ldots\theta}$ when groups $\alpha, \beta, \ldots, \theta$ in the buffer zone are considered as atomistic and $\gamma, \delta, \ldots, K-\theta$ as having a low resolution. The Lagrangian obtained by treating the active and buffer zone at the atomistic resolution can be written as

$$\begin{aligned} L_{\alpha\beta\ldots K} = {} & L + \sum_{\alpha=1,\ldots,K} (L_\alpha - L) \\ & + \sum_{\substack{\alpha=1,\ldots,K-1 \\ \beta=\alpha+1,\ldots,K}} \left\{ L_{\alpha\beta} - \left[L + \sum_{\mu=\alpha,\beta} (L_\mu - L) \right] \right\} \\ & + \sum_{\substack{\alpha=1,\ldots,K-2 \\ \beta=\alpha+1,\ldots,K-1 \\ \gamma=\beta+1,\ldots,K}} \left[L_{\alpha\beta\gamma} - \left(L + \sum_{\mu=\alpha,\beta,\gamma} (L_\mu - L) \right.\right. \\ & \left.\left. + \sum_{(\sigma,\tau)=(\alpha,\beta),(\alpha,\gamma),(\beta,\gamma)} \left\{ L_{\sigma\tau} - \left[L + \sum_{\mu=\alpha,\beta} (L_\mu - L) \right] \right\} \right) \right] + \cdots, \end{aligned} \tag{5}$$

where L corresponds to $\theta = 0$, i.e. only the groups in the active zone have an atomistic resolution. In the AP procedure of Lagrangian functions all high-resolution Lagrangian contributions from groups α in the buffer zone are smoothed according to their radial coordinate r_α (distance between group α and the centre of the active site)

$$\begin{aligned} L^{\mathrm{AP}} = {} & L + \sum_{\alpha=1,\ldots,K} S_\alpha (L_\alpha - L) \\ & + \sum_{\substack{\alpha=1,,\ldots,K-1 \\ \beta=\alpha+1,,\ldots,K}} S_\alpha S_\beta \left\{ L_{\alpha\beta} - \left[L + \sum_{\mu=\alpha,\beta} (L_\mu - L) \right] \right\} \\ & + \sum_{\substack{\alpha=1,,\ldots,K-2 \\ \beta=\alpha+1,,\ldots,K-1 \\ \gamma=\beta+1,,\ldots,K}} S_\alpha S_\beta S_\gamma \left[L_{\alpha\beta\gamma} - \left(L + \sum_{\mu=\alpha,\beta,\gamma} (L_\mu - L) \right.\right. \\ & \left.\left. + \sum_{(\sigma,\tau)=(\alpha,\beta),(\alpha,\gamma),(\beta,\gamma)} \left\{ L_{\sigma\tau} - \left[L + \sum_{\mu=\alpha,\beta} (L_\mu - L) \right] \right\} \right) \right] + \cdots, \end{aligned} \tag{6}$$

where S_α is a fifth-order spline

$$S_\alpha(\Omega_\alpha) = -6\Omega_\alpha^5 + 15\Omega_\alpha^4 - 10\Omega_\alpha^3 + 1 \tag{7}$$

and Ω_α is given by

$$\Omega_\alpha = \frac{r_\alpha - r_{\min}}{r_{\max} - r_{\min}}, \quad \text{for } r_{\min} < r_\alpha < r_{\max}. \tag{8}$$

Different functional forms for the spline functions S_α are possible and our choice is only motivated by its simplicity and property that the mixed-resolution Lagrangian, L^{AP}, becomes a twice differentiable function.

Substituting Equation (3) into Equation (6), we obtain

$$L^{\mathrm{AP}} = T + \sum_{\beta=1,\dots,K} S_\beta (T_\beta - T) - V^{\mathrm{AP}} = T^{\mathrm{AP}} - V^{\mathrm{AP}}$$

$$= \sum_{\alpha=1}^{N} \frac{1}{2} M_\alpha \mathbf{U}_\alpha^2 + \sum_{k=1}^{AM} \frac{1}{2} m_k \mathbf{u}_k^2$$

$$+ \sum_{\beta=1}^{K} \left[(1 - S_\beta) \frac{1}{2} M_\beta \mathbf{U}_\beta^2 + S_\beta \sum_{i=1}^{A} \frac{1}{2} m_i \mathbf{u}_i^2 \right] - V^{\mathrm{AP}}, \quad (9)$$

where T is the kinetic energy of the active zone treated on an all-atomistic level and all other groups treated on a low-resolution level, T_β is the kinetic energy of the active zone and group β in the buffer zone treated on an all-atomistic level and all other groups treated on a low-resolution level, T^{AP} is the mixed-resolution kinetic energy, and V^{AP} is the potential energy calculated with the permuted AP method [36] applied to the mixed atomistic and low-resolution potential energy functions, $V^{(\mu,\mathrm{v})}$, i.e.

$$V^{\mathrm{AP}} = V + \sum_{\alpha=1,\dots,K} S_\alpha (V_\alpha - V)$$

$$+ \sum_{\substack{\alpha=1,\dots,K-1 \\ \beta=\alpha+1,\dots,K}} S_\alpha S_\beta \left\{ V_{\alpha\beta} - \left[V + \sum_{\mu=\alpha,\beta} (V_\mu - V) \right] \right\}$$

$$+ \sum_{\substack{\alpha=1,\dots,K-2 \\ \beta=\alpha+1,\dots,K-1 \\ \gamma=\beta+1,\dots,K}} S_\alpha S_\beta S_\gamma \left[V_{\alpha\beta\gamma} - \left(V + \sum_{\mu=\alpha,\beta,\gamma} (V_\mu - V) \right. \right.$$

$$\left. \left. + \sum_{(\sigma,\tau)=(\alpha,\beta),(\alpha,\gamma),(\beta,\gamma)} \left\{ V_{\sigma\tau} - \left[V + \sum_{\mu=\alpha,\beta} (V_\mu - L) \right] \right\} \right) \right] + \cdots, \quad (10)$$

where the nomenclature is similar to Equation (6). Equation (6) can be simplified to

$$L^{\mathrm{AP}} = \sum_{w=x,y,z} \left\{ \sum_{\alpha=1}^{N} \frac{1}{2} M_\alpha U_{\alpha,w}^2 + \sum_{k=1}^{AM} \frac{1}{2} m_k u_{k,w}^2 \right.$$

$$\left. + \sum_{\beta=1}^{K} \left[(1 - S_\beta) \frac{1}{2} M_\beta \mathbf{U}_\beta^2 + S_\beta \sum_{i=1}^{A} \frac{1}{2} m_i \mathbf{u}_i^2 \right] \right\} - V^{\mathrm{AP}}$$

$$= \sum_{w=x,y,z} \left\{ \sum_{\alpha=1}^{N} \frac{1}{2} M_\alpha U_{\alpha,w}^2 + \sum_{k=1}^{AM} \frac{1}{2} m_k u_{k,w}^2 \right.$$

$$+ \sum_{\beta=1}^{K} \left[(1 - S_\beta) \frac{1}{2} (u_{1,w}^{(\beta)} \dots u_{A,w}^{(\beta)}) \mathbf{M}_D^{(\beta)} \begin{pmatrix} u_{1,w}^{(\beta)} \\ \vdots \\ u_{A,w}^{(\beta)} \end{pmatrix} \right.$$

$$\left. \left. + S_\beta \frac{1}{2} (u_{1,w}^{(\beta)} \dots u_{A,w}^{(\beta)} \mathbf{M}_M^{(\beta)} \begin{pmatrix} u_{1,w}^{(\beta)} \\ \vdots \\ u_{A,w}^{(\beta)} \end{pmatrix} \right] \right\} - V^{\mathrm{AP}}$$

$$= \sum_{w=x,y,z} \left[\sum_{\alpha=1}^{N} \frac{1}{2} M_\alpha U_{\alpha,w}^2 + \sum_{k=1}^{AM} \frac{1}{2} m_k u_{k,w}^2 \right.$$

$$\left. + \sum_{\beta=1 \dots K} \frac{1}{2} (u_{1,w}^{(\beta)} \dots u_{A,w}^{(\beta)} \boldsymbol{\Lambda}^{(\beta)} \begin{pmatrix} u_{1,w}^{(\beta)} \\ \vdots \\ u_{A,w}^{(\beta)} \end{pmatrix} \right] - V^{\mathrm{AP}}, \quad (11)$$

where $\mathbf{u}_k^{(\beta)}$ is the velocity of atom k of group β in the buffer zone, and $\boldsymbol{\Lambda}^{(\beta)}$ is a coordinate-dependent mass matrix of group β. The mass matrix $\boldsymbol{\Lambda}^{(\beta)}$ is given by

$$\boldsymbol{\Lambda}^{(\beta)} = \left[(1 - S_\beta) \mathbf{M}_D^{(\beta)} + S_\beta \mathbf{M}_M^{(\beta)} \right] \quad (12)$$

with

$$\mathbf{M}_D^{(\beta)} \equiv \left[\frac{1}{M_\beta} \begin{pmatrix} m_1^{(\beta)} \\ \vdots \\ m_A^{(\beta)} \end{pmatrix} \left(m_1^{(\beta)} \dots m_A^{(\beta)} \right) \right] \quad (13)$$

and

$$\mathbf{M}_M^{(\beta)} \equiv \begin{bmatrix} m_1^{(\beta)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & m_A^{(\beta)} \end{bmatrix}. \quad (14)$$

Equation (11) shows that the mixed-resolution Lagrangian for the system allows for a change in resolution of selected groups of atoms during a MD simulation by making the kinetic energy part of the adaptively partitioned Lagrangian a functional of the splines. As mentioned in Section 2.1, such a coordinate dependent kinetic energy functional is a requirement for any conservative mixed-resolution system.

2.3 *Adaptive partioning of Hamiltonian functions*

Having defined a mixed-resolution Lagrangian function, we can calculate the generalised momenta $p_i = \partial L / \partial u_i$

and the mixed-resolution Hamiltonian function H^{AP} [35]. We note that this procedure is possible since the phase space dimensionality of the system is constant during an APL simulation and equal to the all-atomistic phase space dimensionality. Only the number of interaction sites and degrees of freedom are reduced and vary during an APL simulation. Consequently, we obtain

$$\begin{aligned}
H^{\mathrm{AP}} = {} & \sum_i u_i p_i - L^{\mathrm{AP}} = H + \sum_{\alpha=1,\dots,K} S_\alpha (H_\alpha - H) \\
& + \sum_{\substack{\alpha=1,\dots,K-1\\ \beta=\alpha+1,\dots,K}} S_\alpha S_\beta \left\{ H_{\alpha\beta} - \left[H + \sum_{\mu=\alpha,\beta} (H_\mu - H) \right] \right\} \\
& + \sum_{\substack{\alpha=1,\dots,K-2\\ \beta=\alpha+1,\dots,K-1\\ \gamma=\beta+1,\dots,K}} S_\alpha S_\beta S_\gamma \Biggl[H_{\alpha\beta\gamma} - \Biggl(H + \sum_{\mu=\alpha,\beta,\gamma} (H_\mu - H) \\
& + \sum_{(\sigma,\tau)=(\alpha,\beta),(\alpha,\gamma),(\beta,\gamma)} \left\{ H_{\sigma\tau} - \left[H + \sum_{\mu=\alpha,\beta} (H_\mu - H) \right] \right\} \Biggr) \Biggr] + \cdots \\
= {} & T^{\mathrm{AP}} + V^{\mathrm{AP}} \\
= {} & \sum_{w=x,y,z} \Biggl[\sum_{\alpha=1}^{N} \frac{P_{\alpha,w}^2}{2M_\alpha} + \sum_{k=1}^{AM} \frac{p_{k,w}^2}{2m_k} \\
& + \sum_{\beta=1\dots K} \frac{1}{2} (p_{1,w}^{(\beta)} \cdots p_{A,w}^{(\beta)}) (\mathbf{\Lambda}^{(\beta)})^{-1} \begin{pmatrix} p_{1,w}^{(\beta)} \\ \vdots \\ p_{A,w}^{(\beta)} \end{pmatrix} \Biggr] + V^{\mathrm{AP}}, \qquad (15)
\end{aligned}$$

where the Hamiltonian description is analogous to Equation (6) and P_α is the total momentum of group α.

The main advantage of having derived a mixed-resolution Hamiltonian is that we can analytically prove that the Hamiltonian conserves energy and linear and angular momentum. Since insertion or removal of atomistic structure on the boundaries of the buffer zone does not alter the kinetic energy, potential energy, or immediate trajectory of the group, the proof that the Hamiltonian conserves energy follows directly from the derivation of the Hamiltonian and the fact that the Hamiltonian has been shown to simply be the sum of the mixed-resolution kinetic and potential energies [35].

To show that the mixed-resolution Hamiltonian, H^{AP}, conserves linear and angular momentum, it is necessary to show that the Poisson bracket of the linear momentum and the Hamiltonian, and the Poisson bracket of the angular momentum and the Hamiltonian are zero [35]. Requiring conservation of linear and angular momentum imposes constraints on the definition of low-resolution sites and Heyden and Truhlar [34] have recently proven that for group coordinates positioned at the centre of mass of their underlying atomistic structure, linear and angular momentum are mathematically conserved.

2.4 Extensions to multiple active zones

In this section, we discuss necessary extensions of the APL method to study systems with multiple active zones, including the possibility that they merge during the simulation. Examples of systems where this extension could be useful are solutions containing multiple ions and counterions, or a system of polymer molecules interacting with a solid surface.

We assume that there are O active zones in the system that are each surrounded by a buffer zone. All active zones and their buffer zones are immersed in a single environmental zone. As before, the buffer and environmental zones contain groups, and each active zone is centred on an active site. Altogether, there are K groups in the buffer zones. The first step of this extended APL method is to calculate the distance $r_{\alpha j}$ of all K groups α in the buffer zones to all O active sites j, i.e. we generalise r_α of previous sections to $r_{\alpha j}$. Assuming spherical active zones, dimensionless distances can be defined as

$$\Omega_{\alpha j} = \frac{r_{\alpha j} - r_{\min,j}}{r_{\max,j} - r_{\min,j}}, \text{ for } r_{\min,j} < r_{\alpha j} < r_{\max,j}, \qquad (16)$$

where $r_{\min,j}$ and $r_{\max,j}$ define the buffer zone around the active zone j. (In principle, the APL method is not limited to spherical active and buffer zones. In fact, both zones can have an arbitrary shape and only the parameter $\Omega_{\alpha j}$, defined in Equation (16), has to be adjusted to the geometric shape of both zones.) Next, the spline functions $\tilde{S}_{\alpha j}$ are calculated

$$\tilde{S}_{\alpha j}(\Omega_{\alpha j}) = -6\Omega_{\alpha j}^5 + 15\Omega_{\alpha j}^4 - 10\Omega_{\alpha j}^3 + 1. \qquad (17)$$

To define a global smoothing function S_α of group α in the system that is zero only if all $\tilde{S}_{\alpha j}$ $(j = 1,\dots,O)$ are zero, one whenever at least one $\tilde{S}_{\alpha j}$ $(j = 1,\dots,O)$ is one, and increases monotonously with increasing $\tilde{S}_{\alpha j}$ $(j = 1,\dots,O)$, we set

$$S_\alpha = 1 - \prod_{j=1}^{O} (1 - \tilde{S}_{\alpha j}). \qquad (18)$$

The only change in this extended APL method is that Equation (18) replaces Equation (7) and L in Equation (6) is now the Lagrangian with all groups present in at least one active zone treated at a high resolution, L_α is the Lagrangian with all groups present in at least one active zone and group α treated at a high resolution, etc. Figure 2 illustrates the use of this extended APL method for the study of a polymer–solid surface system in which selected surface atoms are used to define an active interface zone.

2.5 Extensions to various thermodynamic ensembles

Having defined a single Lagrangian function for a mixed-resolution system, standard procedures for extending the Lagrangian to conditions such as constant temperature and pressure can be applied [2,37–39]. We only stress here that the effective number of degrees of freedom, D, for a specific time or configuration can (as the kinetic energy) be explicitly dependent on coordinates and, furthermore, be a non-integer number

$$D = 3N + 3AM + 3K + 3\sum_{\alpha=1,\ldots,K} S_\alpha(A-1). \tag{19}$$

Properties such as the average temperature, $\langle T\rangle$, can still be calculated as

$$\langle T\rangle = 2\left\langle \frac{T^{\mathrm{AP}}}{D}\right\rangle, \tag{20}$$

where $\langle\cdots\rangle$ denotes a phase space average.

3. An explicit symplectic integrator for mixed-resolution simulations

In addition to being able to define a mixed-resolution Hamiltonian that mathematically conserves energy and linear and angular momentum, the practical applicability of any mixed-resolution technique requires the availability of an efficient symplectic integrator (geometric integrator that numerically conserves energy). Since there exists in general no explicit symplectic integrator for Hamiltonian functions with coordinate-dependent kinetic energies [40], we previously derived an implicit symplectic generalisation of the leapfrog algorithm that solves Hamilton's equations of motion with a single force evaluation per time step [34]. Implicit integrators are usually less convenient to use, more elaborate to implement, and slower. Therefore, we present in this section an explicit symplectic integrator for mixed-resolution systems. This algorithm can be derived from the generalised leapfrog algorithm if we require that all group coordinates (low-resolution sites) be consistently placed on the centre of mass of the underlying atomistic structure [41].

For most high- and low-resolution sites, the algorithm is equivalent to the standard leapfrog algorithm and consists of three steps

$$\begin{aligned}
&(1)\quad p_{k,v}^{(n+(1/2))} = p_{k,v}^{(n)} - \frac{\Delta t}{2}\frac{\partial V^{\mathrm{AP},n}}{\partial q_{k,v}},\\
&(2)\quad q_{k,v}^{(n+1)} = q_{k,v}^{(n)} + \Delta t\frac{p_{k,v}^{(n+(1/2))}}{m_k},\\
&(3)\quad p_{k,v}^{(n+1)} = p_{k,v}^{(n+(1/2))} - \frac{\Delta t}{2}\frac{\partial V^{\mathrm{AP},n+1}}{\partial q_{k,v}},
\end{aligned} \tag{21}$$

where n, $n+(1/2)$ and $n+1$ symbolise a point in time, $v \in \{x,y,z\}, p_{k,v}$ is the generalised momentum of atom k in v-direction, $q_{k,v}$ is the generalised coordinate of atom k in v-direction, and Δt is the time step. Only for groups α in the buffer zone and atoms (or groups) that define the centre of the active zone is it necessary to change the integration algorithm. Step (1) involves now two steps

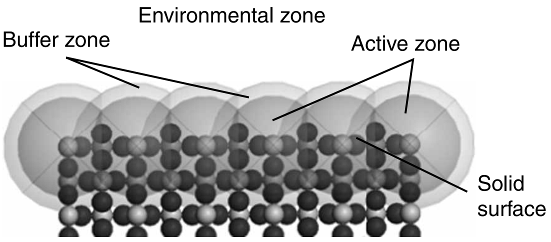

Figure 2. Partitioning of a polymer (not shown)/solid surface system into an active zone specified by multiple surface atoms, an environmental zone, and a buffer zone separating the active from the environmental zone. Groups of atoms in the active zone have a high (atomistic) resolution, groups in the environmental zone have a low (CG) resolution, and groups in the buffer zone have a mixed resolution.

$$\begin{aligned}
&(1\mathrm{a})\quad \tilde{p}_{k,v}^{(n+(1/2))} = p_{k,v}^{(n)} - \frac{\Delta t}{2}\,\frac{\partial V^{\mathrm{AP},n}}{\partial q_{k,v}}\\
&(1\mathrm{b})\quad p_{k,v}^{(n+(1/2))} = \tilde{p}_{k,v}^{(n+(1/2))} + \frac{\Delta t}{4}\sum\nolimits_{w=x,y,z}\sum\nolimits_{\beta=1,\ldots,K}\\
&\left[\left(\tilde{p}_{1,w}^{(\beta)}\cdots\tilde{p}_{A,w}^{(\beta)}\right)^{(n+(1/2))}(\boldsymbol{\Lambda}^{(\beta,n)})^{-1}\right.\\
&\left.\left[\mathbf{M}_M^{(\beta)} - \mathbf{M}_D^{(\beta)}\right](\boldsymbol{\Lambda}^{(\beta,n)})^{-1}\begin{pmatrix}\tilde{p}_{1,w}^{(\beta)}\\ \vdots\\ \tilde{p}_{A,w}^{(\beta)}\end{pmatrix}^{(n+(1/2))}\frac{\partial S_\beta^{(n)}}{\partial q_{k,v}}\right],
\end{aligned} \tag{22}$$

where $\tilde{p}_{k,v}$ has no specific physical meaning. Similarly, step (2) becomes a two-step procedure

$$(2\mathrm{a})\quad \begin{pmatrix}\tilde{q}_{1,v}\\ \vdots\\ \tilde{q}_{A,v}\end{pmatrix}^{(n+1)} = \begin{pmatrix}q_{1,v}\\ \vdots\\ q_{A,v}\end{pmatrix}^{(n)} + \frac{\Delta t}{\mathrm{M}_\alpha}\sum\nolimits_{k=1}^{A} p_{k,v}^{(\alpha)(n+1/2)}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix},$$

$$\text{(2b)}\quad \begin{pmatrix} q_{1,v} \\ \vdots \\ q_{A,v} \end{pmatrix}^{(n+1)} = \begin{pmatrix} q_{1,v} \\ \vdots \\ q_{A,v} \end{pmatrix}^{(n)} + \frac{\Delta t}{2} \left[\begin{matrix} \left(\boldsymbol{\Lambda}^{(\alpha,n)}\right)^{-1} \begin{pmatrix} p_{1,v}^{(\alpha)} \\ \vdots \\ p_{A,v}^{(\alpha)} \end{pmatrix}^{(n+(1/2))} + \\ \left(\tilde{\boldsymbol{\Lambda}}^{(\alpha,n+1)}\right)^{-1} \begin{pmatrix} p_{1,v}^{(\alpha)} \\ \vdots \\ p_{A,v}^{(\alpha)} \end{pmatrix}^{(n+(1/2))} \end{matrix} \right], \tag{23}$$

where we have used vector notation for atoms belonging to a specific group and the coordinates $(\tilde{q}_{1,v}, \ldots, \tilde{q}_{A,v})^{\mathrm{T}}$ are used to calculate the mass matrix $\tilde{\boldsymbol{\Lambda}}^{(\alpha,n+1)}$. We note that steps 2(a) and (b) reduce for atoms defining the centre of the active zone to step (2) in the standard leapfrog algorithm illustrated in Equation (21). Finally, step (3) becomes

$$\begin{aligned} \text{(3)}\ p_{k,v}^{(n+1)} = {} & p_{k,v}^{(n+(1/2))} - \frac{\Delta t}{2}\frac{\partial V^{\mathrm{AP},n+1}}{\partial q_{k,v}} \\ & + \frac{\Delta t}{4}\sum\nolimits_{w=x,y,z}\sum\nolimits_{\beta=1,\ldots,K} \left[\left(p_{1,w}^{(\beta)} \cdots p_{A,w}^{(\beta)}\right)^{(n+(1/2))} \right. \\ & \times \left(\boldsymbol{\Lambda}^{(\beta,n+1)}\right)^{-1} \left[\mathbf{M}_M^{(\beta)} - \mathbf{M}_D^{(\beta)}\right] \left(\boldsymbol{\Lambda}^{(\beta,n+1)}\right)^{-1} \\ & \left. \times \begin{pmatrix} p_{1,w}^{(\beta)} \\ \vdots \\ p_{A,w}^{(\beta)} \end{pmatrix}^{(n+1/2)} \frac{\partial S_\beta^{(n+1)}}{\partial q_{k,v}} \right]. \end{aligned} \tag{24}$$

APL simulations can be performed with an all-atomistic resolution in all zones (the force-field still has different resolutions). In this case, the Hamiltonian propagates the atomistic structure of groups in the environmental zone in parallel to the group coordinates, and the APL scheme and integrator are time-reversible. The main disadvantage of this approach is that the whole all-atomistic system, including the atomistic structure of groups in the environmental zone, has to be equilibrated with an all-atomistic force field to obtain a statistically unbiased initial structure. This step is often computationally unfeasible. To eliminate this extra procedure, we propose to use the APL algorithm with reduced-dimensionality in the environmental zone, integrate only the group coordinates of groups in the environmental zone, and only insert the atomistic structure of a group as it enters the buffer zone from the environmental zone. Thus, we need an algorithm for determining atomistic structure of a group during the simulation. There are two requirements: (i) the procedure should generate appropriate atomic coordinates for the simulation conditions without changing the group coordinate. (ii) The configuration space should be sampled with a probability similar to that of an all-atomistic simulation. One possible procedure is to randomly generate an atomistic configuration for the group entering the buffer zone that satisfies requirement (i), and then to perform a constrained MD simulation to relax the newly placed atoms in the force field of all the other atoms and groups. During this process, only the newly placed atoms move (the group coordinates are fixed) and all atoms in the buffer and core zone are considered explicit. Other procedures for inserting atomistic structure based on Monte Carlo methods are also possible [42,43].

We note that no velocities but only atomic coordinates have to be generated in the APL scheme when a group enters the buffer zone from the environmental zone. For groups on the boundary between the buffer and environmental zone, the atomic velocities are the same as the velocity of the group coordinate. The Hamiltonian automatically reintroduces intra-group velocities as a group in the buffer zone moves towards the active zone and it removes intra-group kinetic energy as a group in the buffer zone moves towards the environmental zone.[2] Furthermore, the AP potential energy is independent of the atomistic structure of a group on the boundary between buffer and environmental zone so that the insertion of atomic structure yields both a continuous kinetic energy and a continuous potential energy, and the APL scheme is time-reversible in group coordinates and momenta.

4. The main challenge in mixed-resolution multiscale simulations

The APL scheme is currently the only algorithm that solves the Euler–Lagrange equations of motion for a conservative mixed-resolution system, is computationally efficient and can easily be extended to any thermodynamic ensemble. Nevertheless, the applicability of the APL scheme will greatly depend on our ability to define a physically meaningful mixed-resolution Lagrangian/Hamiltonian. The problem is in many respects similar to those encountered in developing physical CG potential energy functions. Again, we require that a good mixed-resolution Hamiltonian satisfies

$$\begin{aligned}\exp(-F/k_{\mathrm{B}}T) &= \frac{1}{N_{\mathrm{AA}}!h^{3N_{\mathrm{AA}}}}\int \mathrm{d}\mathbf{q}_{\mathrm{AA}}\,\mathrm{d}\mathbf{p}_{\mathrm{AA}} \\ &\quad\times \exp(-H_{\mathrm{AA}}(\mathbf{q}_{\mathrm{AA}},\mathbf{p}_{\mathrm{AA}})/k_{\mathrm{B}}T) \\ &\approx \frac{1}{N_{\mathrm{AA}}!h^{3N_{\mathrm{AA}}}}\int \mathrm{d}\mathbf{q}_{\mathrm{AA}}\,\mathrm{d}\mathbf{p}_{\mathrm{AA}} \\ &\quad\times \exp(-H^{\mathrm{AP}}(\mathbf{q}_{\mathrm{AA}},\mathbf{p}_{\mathrm{AA}})/k_{\mathrm{B}}T),\end{aligned} \tag{25}$$

where N_{AA} is the total number of atoms in the system, h is the Planck's constant, $H_{\mathrm{AA}}(\mathbf{q}_{\mathrm{AA}}, \mathbf{p}_{\mathrm{AA}})$ is the atomistic Hamiltonian, $\mathbf{q}_{\mathrm{AA}}$ and $\mathbf{p}_{\mathrm{AA}}$ are the atomistic phase space variables, and only the domain over which the integration is performed is different in the AA and mixed-resolution cases. Consequently, the mixed-resolution Hamiltonian or more precisely, the mixed-resolution potential energy, V^{AP}, is a free energy that includes some entropic contributions. If we assume that the AP procedure (the mixing of potential energy functions) is not a critical step that introduces unphysical behaviour, then Equation (25) is equivalent to requiring for all potential energy functions, $V^{(\mu,v)}$, with a fixed number of v CG and μ atomistic groups to satisfy

$$\begin{aligned}\exp(-F/k_{\mathrm{B}}T) &= (\text{const.})\int \mathrm{d}\mathbf{q}_{\mathrm{AA}}\exp(-V_{\mathrm{AA}}(\mathbf{q}_{\mathrm{AA}})/k_{\mathrm{B}}T) \\ &\approx (\text{const.}^{(\mu,v)})\int \mathrm{d}\mathbf{q}^{\mu}_{\mathrm{AA}}\,\mathrm{d}\mathbf{q}^{v}_{\mathrm{CG}} \\ &\quad\times \exp(-V^{(\mu,v)}(\mathbf{q}^{\mu}_{\mathrm{AA}},\mathbf{q}^{v}_{\mathrm{CG}})/k_{\mathrm{B}}T),\end{aligned} \tag{26}$$

where $\mathbf{q}^{\mu}_{\mathrm{AA}}$ and $\mathbf{q}^{v}_{\mathrm{CG}}$ are the atomistic and CG coordinates, respectively, and (const.) and (const.$^{(\mu,v)}$) are normalisation constants. We note that satisfying Equations (25) and (26) is more elaborate than satisfying Equation (1) since we are also required to determine the state dependent constants, (const.$^{(\mu,v)}$). In other words, all potential energy functions, $V^{(\mu,v)}$, have to be determined with reference to the same point or an average of points in the all-atomistic phase space of the system (e.g. a point in configuration space with all momenta equal to zero), and the absolute magnitude of the potential energy functions effects Equation (25) and is thermodynamic-state dependent. Future work will show how much more difficult the determination of a physical mixed-resolution Hamiltonian, H^{AP}, is in comparison to obtaining a physical CG force field.

Finally, we note that for K groups in the buffer zone, 2^K potential energy functions have to be evaluated to determine the AP potential energy, V^{AP}. While this step can computationally become very time consuming, we note that for classical mixed-resolution force-fields that consist of additive many-body interaction terms, the computational overhead is usually negligible. For example, for pairwise potentials, Equation (10) reduces to

$$\begin{aligned}V^{\mathrm{AP}} &= V + \sum_{\alpha=1,\ldots,K} S_{\alpha}(V_{\alpha} - V) \\ &\quad+ \sum_{\substack{\alpha=1,\ldots,K-1\\ \beta=\alpha+1,\ldots,K}} S_{\alpha}S_{\beta}\left\{V_{\alpha\beta} - \left[V + \sum_{\mu=\alpha,\beta}\left(V_{\mu} - V\right)\right]\right\}\end{aligned} \tag{27}$$

and the computational effort scales as $\{(1/2)(AM{+}N{+}K)(AM{+}N{+}K{-}1){+}(1/2)AK(AM{+}N){+}(1/2)AK(AK{-}1)\}$. Consequently, the computational overhead of the APL procedure, $\{(1/2)AK(AM + N) + (1/2)AK(AK - 1)\}$, is negligible if the number of atoms in the buffer zone, AK, is small in comparison to the number of interaction sites in the active and environmental zone, $AM + N$.

5. Mixed-resolution MD simulations of supercritical methane

In this section, we numerically test the energy and linear momentum conservation properties of the APL method and the symplectic integrator. The model system consists of 64 methane molecules in a periodic box with a box length of 20.8 Å. (The density of methane is 0.189 g/cm^3 and therefore slightly above the critical density of methane which is 0.162 g/cm^3.) The centre of mass of one methane molecule is chosen to be the centre of the active zone in which the methane molecules have atomistic resolution. The radius of the active zone is $r_{\min} = 7$ Å, the buffer zone is 1 Å thick, and a single methane molecule in the environmental zone is CG into a single low-resolution site. For the case considered here with no covalent bonds between groups, there is considerable freedom in the choice of the mixed-resolution potential energy functional for a fixed number of v CG and μ atomistic groups, but one reasonable possibility is

$$V^{(\mu,v)} = V^{\mu+v} + (V^{\mu}_{\mathrm{AA}} - V^{\mu}) + 3(A-1)vk_{\mathrm{B}}T_0, \tag{28}$$

where $V^{\mu+v}$ is the potential energy of the entire system described by a low-resolution force field, V^{μ}_{AA} is the potential energy of the μ active groups described by an atomistic force field, V^{μ} is the potential energy of the μ active groups described by the low-resolution force field, and A is the number of atoms per group, i.e. $A = 5$. We have used the force-shifted CHARMM potential [44] with a cut-off radius of 10 Å to describe the atomistic methane molecules and the force-shifted OPLS-UA potential [45] with a cut-off radius of 10 Å to describe the CG methane molecules. The Lorentz–Berthelot mixing rule [1] has been used for calculating the interaction between high- and low-resolution sites. Since the low-resolution force field does not include energy contributions of the implicit degrees of freedom at a finite temperature, Equation (28) includes k_BT_0 for each implicit degree of freedom in the low-resolution description of the potential energy $V^{(\mu,v)}$, where $T_0 = 300\,\text{K}$ is the average simulation temperature we initially intended to study (the actual average temperature of our simulation in the NVE ensemble has been 319 K). We note that the choice of k_BT_0 per degree of freedom is motivated by considering all implicit degrees of freedom to be harmonic [46]. While the described procedure for developing a mixed-resolution potential energy function for methane molecules does not exactly fulfil Equation (26), we note that the objective of the performed simulations has not been to realistically model supercritical methane but to test the conservation properties of our algorithm. Future work will show how difficult it is to determine a mixed-resolution Hamiltonian that permits the simulation of complex physical phenomena. Figure 3 illustrates the kinetic, potential and total energy of the system during a simulation in the NVE ensemble. For an integration time step of 0.05 fs and a simulation length of 1 ns, the total energy is conserved during the simulation. (We note that the integration time step has not been optimised for efficiency and the computational efficiency can be significantly improved if multiple-time-step techniques are used [47].) Figure 4, furthermore, shows that the linear momentum of the system is also approximately conserved during the simulation. While there seems to be a small drift in the total linear momentum per methane molecule, the size of the drift $\sim 10^{-14}$ amu Å ns/fs is negligible. Overall, the developed explicit symplectic integrator of our APL method illustrates excellent numerical conservation properties.

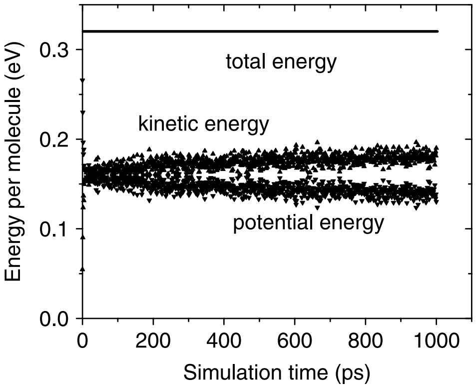

Figure 3. Kinetic, potential, and total energy per methane molecule during a microcanonical (NVE) simulation of 64 CH_4 molecules in a periodic box (box length = 20.8 Å) and at an average temperature of 319 K. The centre of mass of a specific CH_4 molecule defines the centre of a high resolution zone ($r_{\min} = 7$ Å ; $r_{\max} = 8$ Å).

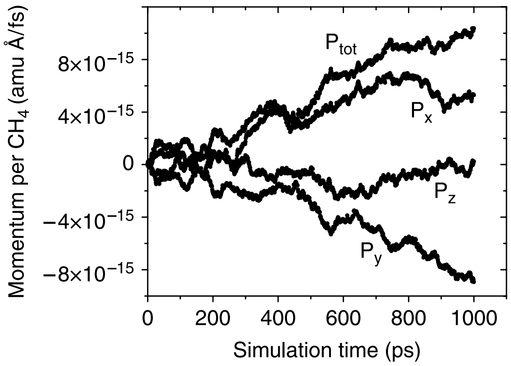

Figure 4. Linear momentum per CH_4 molecule during a microcanonical (NVE) simulation of 64 CH_4 molecules in a periodic box (box length = 20.8 Å) and at an average temperature of 319 K. The centre of mass of a specific CH_4 molecule defines the centre of a high resolution zone ($r_{\min} = 7$ Å ; $r_{\max} = 8$ Å).

6. Conclusions

In conclusion, we have derived a Hamiltonian and presented an explicit symplectic integrator for conservative mixed-resolution systems that allows for a change in resolution of selected groups of atoms during a MD simulation. We numerically verified our previous analytical proof that a NVE simulation protocol with this Hamiltonian and numerical integrator conserves energy and momentum. Extensions of our novel method to various thermodynamic ensembles and multiple active zones that might merge during a simulation have also been described. Overall, for systems that require atomistic resolution only for a small part or parts of the simulation system, the APL method permits a simulation with accuracy comparable to an atomistic one at the computational cost of a CG one. The main challenge in mixed-resolution simulations is the determination of a physically meaningful mixed-resolution force field. Both 'inverse' coarse-graining techniques [48,49] (wherein macroscopic thermodynamic properties are used to parameterise the mixed-resolution force field) and multiscale coarse-graining methods [5,10] (that use atomistic molecular simulations in a 'bottom up' approach to parameterise mixed-resolution force fields) will likely play a major role in their development. Considering that mixed-resolution simulations have the potential to be more predictive than CG simulations, the presented multiscale approach together with the efficient generation of accurate mixed atomistic and CG force fields provide a new route to model complex systems that require only locally (i.e. close to an active site or interface) an atomistic level of theory.

Acknowledgements

The authors are grateful to Donald G. Truhlar for helpful discussions that led to presented work. Furthermore, we gratefully acknowledge the Minnesota Supercomputing Institute for computer time.

Notes

1. The curl of a force field is zero if the force field is conservative, but it can be shown that the curl of the force field in the AdResS scheme is in general not equal to zero for groups in the buffer zone.
2. Reference [34] erroneously mentions that the group coordinate should be placed on one of the atoms of the group. Such a procedure prevents the Hamiltonian from reintroducing the intra-group velocities and is therefore not recommended.

References

[1] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1989.

[2] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed., Academic Press, London, 2002.

[3] F. Mueller-Plathe, *Coarse-graining in polymer simulation: from the atomistic to the mesoscopic scale and back*, ChemPhysChem 3 (2002), pp. 754–769.

[4] S.H.L. Klapp, D.J. Diestler, and M. Schoen, *Why are effective potentials 'soft'?*, J. Phys. Condens. Matter 16 (2004), pp. 7331–7352.

[5] S. Izvekov and G.A. Voth, *Multiscale coarse graining of liquid-state systems*, J. Chem. Phys. 123 (2005), 134105.

[6] C.F. Lopez, P.B. Moore, J.C. Shelley, M.Y. Shelley, and M.L. Klein, *Computer simulation studies of biomembranes using a coarse grain model*, Comput. Phys. Commun. 147 (2002), pp. 1–6.

[7] P.K. Depa and J.K. Maranas, *Speed up of dynamic observables in coarse-grained molecular-dynamics simulations of unentangled polymers*, J. Chem. Phys. 123 (2005), 094901.

[8] M. Muller, K. Katsov, and M. Schick, *Biological and synthetic membranes: what can be learned from a coarse-grained description?*, Phys. Rep. 434 (2006), pp. 113–176.

[9] G.A. Voth (ed.), *Coarse-graining of Condensed Phase and Biomolecular Systems*, CRC Press, Boca Raton, FL, 2009.

[10] S. Izvekov and G.A. Voth, *A multiscale coarse-graining method for biomolecular systems*, J. Phys. Chem. B 109 (2005), pp. 2469–2473.

[11] W.G. Noid, J.W. Chu, G.S. Ayton, V. Krishna, S. Izvekov, G.A. Voth, A. Das, and H.C. Andersen, *The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models*, J. Chem. Phys. 128 (2008), 244114.

[12] W.G. Noid, P. Liu, Y. Wang, J.W. Chu, G.S. Ayton, S. Izvekov, H.C. Andersen, and G.A. Voth, *The multiscale coarse-graining method. II. Numerical implementation for coarse-grained molecular models*, J. Chem. Phys. 128 (2008), 244115.

[13] L. Delle Site, C.F. Abrams, A. Alavi, and K. Kremer, *Polymers near metal surfaces: selective adsorption and global conformations*, Phys. Rev. Lett. 89 (2002), 156103.

[14] E. Villa, A. Balaeff, L. Mahadevan, and K. Schulten, *Multiscale method for simulating protein-DNA complexes*, Multiscale Model. Simul. 2 (2004), pp. 527–553.

[15] M. Neri, C. Anselmi, M. Cascella, A. Maritan, and P. Carloni, *Coarse-grained model of proteins incorporating atomistic detail of the active site*, Phys. Rev. Lett. 95 (2005), 218102.

[16] C.F. Abrams, *Concurrent dual-resolution Monte Carlo simulation of liquid methane*, J. Chem. Phys. 123 (2005), 234101.

[17] M. Praprotnik, L. Delle Site, and K. Kremer, *Adaptive resolution molecular-dynamics simulation: changing the degrees of freedom on the fly*, J. Chem. Phys. 123 (2005), 224106.

[18] L.M. Dupuy, E.B. Tadmor, R.E. Miller, and R. Phillips, *Finite-temperature quasicontinuum: molecular dynamics without all the atoms*, Phys. Rev. Lett. 95 (2005), 060202.

[19] R.E. Rudd and J.Q. Broughton, *Coarse-grained molecular dynamics: nonlinear finite elements and finite temperature*, Phys. Rev. B 72 (2005), 144104.

[20] D.J. Diestler, H. Zhou, R. Feng, and X.C. Zeng, *Hybrid atomistic-coarse-grained treatment of multiscale processes in heterogeneous materials: a self-consistent-field approach*, J. Chem. Phys. 125 (2006), 064705.

[21] M. Praprotnik, L. Delle Site, and K. Kremer, *Adaptive resolution scheme for efficient hybrid atomistic-mesoscale molecular dynamics simulations of dense liquids*, Phys. Rev. E 73 (2006), 066701.

[22] M. Praprotnik, K. Kremer, and L. Delle Site, *Adaptive molecular resolution via a continuous change of the phase space dimensionality*, Phys. Rev. E 75 (2007), 017701.

[23] M. Praprotnik, K. Kremer, and L. Delle Site, *Fractional dimensions of phase space variables: a tool for varying the degrees of freedom of a system in a multiscale treatment*, J. Phys. A: Math. Theor. 40 (2007), pp. F281–F288.

[24] M. Praprotnik, L. Delle Site, and K. Kremer, *A macromolecule in a solvent: adaptive resolution molecular dynamics simulation*, J. Chem. Phys. 126 (2007), 134902.

[25] M. Praprotnik, S. Matysiak, L. Delle Site, K. Kremer, and C. Clementi, *Adaptive resolution simulation of liquid water*, J. Phys. Condens. Matter 19 (2007), 292201.

[26] B. Ensing, S.O. Nielsen, P.B. Moore, M.L. Klein, and M. Parrinello, *Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics*, J. Chem. Theory Comput. 3 (2007), pp. 1100–1105.

[27] M. Cascella, M.A. Neri, P. Carloni, and M. Dal Peraro, *Topologically based multipolar reconstruction of electrostatic interactions in multiscale simulations of proteins*, J. Chem. Theory Comput. 4 (2008), pp. 1378–1385.

[28] M. Praprotnik, C. Junghans, L. Delle Site, and K. Kremer, *Simulation approaches to soft matter: generic statistical properties vs. chemical details*, Comput. Phys. Commun. 179 (2008), pp. 51–60.

[29] M. Praprotnik, L. Delle Site, and K. Kremer, *Multiscale simulation of soft matter: from scale bridging to adaptive resolution*, Annu. Rev. Phys. Chem. 59 (2008), pp. 545–571.

[30] R. Delgado-Buscalioni, K. Kremer, and M. Praprotnik, *Concurrent triple-scale simulation of molecular liquids*, J. Chem. Phys. (2008), 114110.

[31] S. Matysiak, C. Clementi, M. Praprotnik, K. Kremer, and L. Delle Site, *Modeling diffusive dynamics in adaptive resolution simulation of liquid water*, J. Chem. Phys. (2008), 024503.

[32] C. Junghans, M. Praprotnik, and K. Kremer, *Transport properties controlled by a thermostat: an extended dissipative particle dynamics thermostat*, Soft Matter 4 (2008), pp. 156–161.

[33] M. Neri, M. Baaden, V. Carnevale, C. Anselmi, A. Maritan, and P. Carloni, *Microseconds dynamics simulations of the outer-membrane protease T*, Biophys. J. 94 (2008), pp. 71–78.

[34] A. Heyden and D.G. Truhlar, *Conservative algorithm for an adaptive change of resolution in mixed atomistic/coarse-grained multiscale simulations*, J. Chem. Theory Comput. 4 (2008), pp. 217–221.

[35] H. Goldstein, C.P. Poole, and J.L. Safko, *Classical Mechanics*, 3rd ed., Addison Wesley, San Francisco, CA, 2002.

[36] A. Heyden, H. Lin, and D.G. Truhlar, *Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations*, J. Phys. Chem. B 111 (2007), pp. 2231–2241.

[37] M.E. Tuckerman, C.J. Mundy, and G.J. Martyna, *On the classical statistical mechanics of non-Hamiltonian systems*, Europhys. Lett. 45 (1999), pp. 149–155.

[38] M.E. Tuckerman and G.J. Martyna, *Understanding modern molecular dynamics: techniques and applications*, J. Phys. Chem. B 104 (2000), pp. 159–178.

[39] M.E. Tuckerman, Y. Liu, G. Ciccotti, and G.J. Martyna, *Non-Hamiltonian molecular dynamics: generalizing Hamiltonian phase space principles to non-Hamiltonian systems*, J. Chem. Phys. 115 (2001), pp. 1678–1702.

[40] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge University Press, New York, 2004.

[41] J.H. Park and A. Heyden, J. Chem. Phys., submitted for publication.

[42] G.H. Peslherbe, H.B. Wang, and W.L. Hase, *Monte Carlo sampling for classical trajectory simulations*, Adv. Chem. Phys. 105 (1999), pp. 171–201.

[43] J.J. DePablo and F.A. Escobedo, *Monte Carlo methods for polymeric systems*, Adv. Chem. Phys. 105 (1999), pp. 337–367.

[44] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *Charmm - a program for macromolecular energy, minimization, and dynamics calculations*, J. Comput. Chem. 4 (1983), pp. 187–217.

[45] W.L. Jorgensen, D.S. Maxwell, and Tirado-Rives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids*, J. Am. Chem. Soc. 118 (1996), pp. 11225–11236.

[46] K. Huang, *Statistical Mechanics*, Wiley, New York, 1963.

[47] G.J. Martyna, M.E. Tuckerman, D.J. Tobias, and M.L. Klein, *Explicit reversible integrators for extended systems dynamics*, Mol. Phys. 87 (1996), pp. 1117–1157.

[48] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, and A.H. de Vries, *The MARTINI force field: coarse grained model for biomolecular simulations*, J. Phys. Chem. B 111 (2007), pp. 7812–7824.

[49] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, and Marrink, *The MARTINI coarse-grained force field: extension to proteins*, J. Chem. Theory Comput. 4 (2008), pp. 819–834.