

Metadata of the chapter that will be visualized online

ChapterTitle	Multiscale Molecular Dynamics and the Reverse Mapping Problem	
Chapter Sub-Title		
Chapter CopyRight - Year	Springer Science+Business Media B.V. 2009 (This will be the copyright line in the final PDF)	
Book Name	Trends in Computational Nanomechanics	
Corresponding Author	Family Name	Ensing
	Particle	
	Given Name	Bernd
	Suffix	
	Division	Van't Hoff Institute for Molecular Sciences
	Organization	University of Amsterdam
	Address	Nieuwe Achtergracht 166, 1018, WV, Amsterdam, The Netherlands
	Email	b.ensing@uva.nl

Author	Family Name	Nielsen
	Particle	
	Given Name	Steven O.
	Suffix	
	Division	Department of Chemistry
	Organization	University of Texas at Dallas
	Address	800 West Campbell Road, 75080, Richardson, TX, USA
	Email	steven.nielsen@utdallas.edu

Abstract	<p>Multiscale techniques are becoming increasingly important for molecular simulation as a result of interest in increasingly complex problems involving events occurring over multiple time and length scales. Here, inspired by the success of the multiscale quantum mechanics / molecular mechanics (QM/MM) methods, we introduce a hybrid, adaptive resolution, multiscale molecular dynamics method that combines accurate, atomistic, modeling of key regions of the system with a course-grained modeling of the remainder of the system. Hybrid multiscale methods must solve the interfacial hand-shaking problem of coupling together different levels of description in different spatial regions of the system; in addition, to implement an adaptive resolution algorithm to correctly model diffusive systems, one must have a procedure in place to dynamically change the representation of a molecule, either from a finer to a coarser level or vice versa. We propose a solution to these problems through a detailed energy analysis and the use of a rotational dynamics to align molecular fragments. The algorithms we propose significantly advance the state-of-the-art and should serve to spur significant advances in our ability to model complex chemical systems</p>	
Keywords (separated by '-')	Multiscale - Adaptive resolution - Coarse-graining - Molecular dynamics - Reverse mapping - Rotational dynamics	

CHAPTER 2

MULTISCALE MOLECULAR DYNAMICS AND THE REVERSE MAPPING PROBLEM

BERND ENSING¹ AND STEVEN O. NIELSEN²

¹ *Van't Hoff Institute for Molecular Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands, e-mail: b.ensing@uva.nl*

² *Department of Chemistry, University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080, USA, e-mail: steven.nielsen@utdallas.edu*

Abstract: Multiscale techniques are becoming increasingly important for molecular simulation as a result of interest in increasingly complex problems involving events occurring over multiple time and length scales. Here, inspired by the success of the multiscale quantum mechanics / molecular mechanics (QM/MM) methods, we introduce a hybrid, adaptive resolution, multiscale molecular dynamics method that combines accurate, atomistic, modeling of key regions of the system with a course-grained modeling of the remainder of the system. Hybrid multiscale methods must solve the interfacial hand-shaking problem of coupling together different levels of description in different spatial regions of the system; in addition, to implement an adaptive resolution algorithm to correctly model diffusive systems, one must have a procedure in place to dynamically change the representation of a molecule, either from a finer to a coarser level or vice versa. We propose a solution to these problems through a detailed energy analysis and the use of a rotational dynamics to align molecular fragments. The algorithms we propose significantly advance the state-of-the-art and should serve to spur significant advances in our ability to model complex chemical systems

Keywords: Multiscale, Adaptive resolution, Coarse-graining, Molecular dynamics, Reverse mapping, Rotational dynamics

2.1. INTRODUCTION

In numerical analysis and computer simulations, multiscale techniques are used where possible to obtain higher accuracies for lower computational cost. Weather forecast simulations and fluid dynamics calculations, for example, often make use of solving Navier-Stokes-like differential equations on discrete grids with a variable length spacing, so-called multigrids. The choice of the length scale is arbitrary; perhaps a particular part of the calculation is of greater interest and is therefore performed at a higher resolution or accuracy than the remainder. But more often, the

46 length scales used reflect those of the underlying physics of the simulated process.
47 In the case of weather prediction, mountainous regions with rapidly varying topog-
48 raphy require pressure and temperature evaluations on a denser grid than do flat
49 regions.

50 The properties of materials are governed by processes that take place over a vast
51 range of length and time scales. Creep in polymers and glasses is a very slow pro-
52 cess that can easily have relaxation constants on the order of reciprocal years. On
53 the other hand, the atomic motions in the same materials take place on the fem-
54 tosecond time scale. The electronic motions at the onset of a chemical reaction are
55 even faster, while the chemical reaction itself might take place only a few times
56 per second within a certain amount of material, resulting in a time scale ratio of
57 15 orders of magnitude. This inherent multiscale character of phenomena in mate-
58 rials is seen also in the length scales. A charge transfer or proton transfer chemical
59 reaction can be a very localized process, taking place within a region of radius ten
60 Å ngstroms. The same reaction taking place in the active site of an enzyme involves
61 the catalytic effect of the tens of nanometers-sized protein environment, while in
62 general biological processes taking place in the cell, such as signal transduction
63 or gene expression, are often intricately governed by long-range changes in the
64 environment.

65 For the modeling of molecular phenomena with such inherent multi scale charac-
66 ter, new developments have been made to extend existing simulation techniques that
67 could otherwise only be applied to rather limited ranges of application. Accurate
68 quantum mechanical *ab initio* methods allow for electronic structure calculations
69 using a large supercomputer involving tens to hundreds of atoms for tens of picosec-
70 onds. Neglecting electronic structure, the behavior of molecular systems of up to a
71 million atoms can be simulated for hundreds of nanoseconds using classical molec-
72 ular dynamics techniques. Modeling of even larger systems or of processes that take
73 place on even longer time scales requires one to relinquish an atomistic representa-
74 tion and simulate the motion of effective particles that each comprise several atoms;
75 this is the domain of coarse-grained and mesoscale methods.

76 However, there is another choice: one can use multiscale techniques. For example
77 in quantum chemistry, one can include the extended environment of an electronic
78 process using embedding techniques [1, 2], continuum models [3–5] or hybrid
79 quantum mechanics and classical forcefield (QM/MM) techniques [6]. In classical
80 molecular dynamics simulations implicit solvent models can be employed to reduce
81 the computation cost (with a concomitant loss of accuracy) of including for exam-
82 ple a protein environment. And more recently, so-called hybrid multiscale methods
83 have been developed that mix an atomistic molecular dynamics treatment of part of
84 a system with a lower resolution treatment of the rest [7–15].

85 Popularity of the term *multiscale* has led the word to be used in various con-
86 texts and with an increasing (multi-) scale of meanings. Here, we will distinguish
87 between two types of simulation techniques termed multiscale, namely (1) those
88 in which methods of different accuracy and scale are used sequentially [16, 17],
89 and (2) those in which methods of different accuracy and scale are used simul-
90 taneously [7–12, 18–21]. The first type makes use of a relatively accurate, high

Multiscale Molecular Dynamics and the Reverse Mapping Problem

91 resolution, method to parameterize the less accurate, low resolution, method which
92 is then used for the actual simulation. In a sense, this type of multiscale method
93 includes practically all semi-empirical and empirical methods, such as MD simula-
94 tions that employ forcefields that were first parameterized using quantum chemical
95 (e.g. Hartree-Fock) calculations. Instead, the latter type of multiscale techniques are
96 based on treating different parts of a system with different resolutions and include
97 for example multi-grid methods, multi-timestep techniques, certain Hamiltonian
98 switch or replica exchange methods and hybrid methods. The multiscale method
99 that is the subject of this book chapter is a hybrid method that allows key parts of
100 the system to be treated at a high, atomistic, level of resolution while the rest of the
101 system is modeled at a lower, coarse-grained, level of resolution [7].

102 Hybrid multiscale methods must solve the interfacial hand-shaking problem
103 of coupling together different levels of description in different spatial regions
104 of the system. If the shape or position of these spatial regions is changing in
105 time, or if particles are allowed to move between the spatial regions, a special
106 mechanism must be introduced into the multiscale method to allow particles to
107 dynamically adapt their representation. The “Learn On The Fly” method is an
108 example of such an adaptive hybrid method that has been successfully applied
109 to model the propagation of a crack in a brittle solid, in which only the atoms
110 in the advancing crack tip region are modeled at the QM tight-binding level of
111 theory and the rest are treated with a classical forcefield representation [22]. In
112 this case, a one-to-one mapping exists between the atoms in the quantum rep-
113 resentation and those in the classical representation (i.e. only the number of
114 electronic degrees of freedom differs). Bridging between an atomistic represen-
115 tation and a coarse-grained representation, in which each coarse-grained particle
116 comprises several atoms, however, raises the additional difficulty that the map-
117 ping between the two representations is no longer trivial. In particular, the so-called
118 *reverse* mapping of moving from the lower, coarse-grained, resolution to the higher,
119 atomistic, resolution is fraught with difficulty as it requires the generation of
120 information.

121 The need for adaptive boundaries within multiscale modeling methods becomes
122 more urgent as we move from hybrid high-level QM/low-level QM methods
123 and QM/MM methods, along the accessible time and length scales, to hybrid
124 atomistic/coarse-grained methods. Whereas often static boundary suffice on the
125 relatively small, tens of picoseconds, timescale accessible to QM/MM, as shown
126 for example by its success in modeling enzymatic chemistry, instead adaptive
127 boundaries become crucial on the microsecond timescale domain of coarse-grained
128 models, and therefore also hybrid atomistic/coarse-grained methods, which are
129 particularly developed to study diffusive behavior in complex fluids, such as poly-
130 mers, proteins, and amphiphilic assemblies. Adaptive boundaries are thus expected
131 to be essential in hybrid atomistic/coarse-grained studies of a wide range of
132 applications, such as protein folding, protein-protein interaction, permeation in
133 (trans-membrane) ion-channels and pores, diffusion and adsorption on surfaces and
134 open-framework materials, permeation through polymer electrolytes, self-assembly
135 of nano-materials, etcetera, etcetera.

In the remainder of this book chapter, we describe the construction of a hybrid multiscale molecular dynamics method that bridges regions of high (atomistic) resolution and regions of low (coarse-grained) resolution, paying special attention to the reverse mapping problem that needs to be overcome in order to make the method adaptive. But first we give a brief introduction to atomistic and coarse-grained molecular dynamics and on the forward and reverse mapping between these representations.

2.1.1. Atomistic and Coarse-Grained Molecular Dynamics

Atomistic and coarse-grained molecular dynamics simulations are particle-based methods in which conservative forces, and also sometimes dissipative and random forces, are used to evolve the particles in time; this time evolution is called a trajectory. Such classical trajectories in molecular systems can be computed on present day computers for systems containing $10^5 - 10^6$ particles for as long as about 10^8 discrete time steps. If these particles are chosen to be the atoms, the requirement to accurately sample the molecular vibrations will set the maximum total simulation time to be less than a microsecond. The most cpu-intensive part of the computation is the evaluation of the non-bonded van der Waals and electrostatic interactions, which are typically approximated by pairwise additive 2-body potentials, yielding in principle a quadratic scaling of the computational effort with system size.

For certain long-time or large length scale phenomena, the fastest molecular vibrations are irrelevant, making it desirable to average out these high frequency fluctuations a priori and model directly the representative (coarse grained) particles that move on the mesoscopic length and time scale. To model these large scale motions in complex fluids such as polymers, colloids, surfactants and bio-molecular assemblies, mesoscopic simulation methods have been developed: for instance Dissipative Particle Dynamics [23], Langevin Dynamics [24], and Brownian Dynamics [25], in which the local atomic rattling is simplified to random noise and dissipation terms. The “potential energy surface” on which these coarse-grained particles move can in principle be computed from the high-resolution representation of the system by integrating over all irrelevant fast-frequency degrees of freedom r :

$$U(R) = -k_B T \ln \int dr e^{-V(R,r)/k_B T} \quad (2-1)$$

with k_B Boltzmann’s constant, T the absolute temperature, and V the potential that governs the fundamental interactions in the system. The resulting effective potential felt by the coarse-grained particles, $U(R)$, is actually a free energy surface that is also a function of the thermodynamic variables that define the state at which $U(R)$ is evaluated, for example the temperature and the pressure in an isothermal, isobaric ensemble. Free energy methods, such as the metadynamics method [26, 27], allow for the evaluation of a free energy landscape of a (very) small set of slow variables, which can then be used in a Langevin dynamics to explore the kinetics. However, for the construction of a (many-) particle based coarse-grained model, calculation

Multiscale Molecular Dynamics and the Reverse Mapping Problem

181 of the exact potential $U(R)$ is not feasible in practise and people have devised other
182 approximate methods to build coarse-grained potentials, or forcefields.

183 The level of coarse-graining applied here in the context of hybrid multi-
184 scale molecular dynamics is in between atomistic and mesoscopic. At this level,
185 coarse-grained particles represent approximately the chemical functional groups
186 of molecules, containing in the order of ten atoms each [28]. That way, the
187 same machinery as used for atomistic molecular dynamics can be used for the
188 coarse-grained molecular dynamics, including for example harmonic bond and
189 bend functions and non-bonded van der Waals-like and electrostatic interactions
190 that make up the forcefield. Critically, this lets us use existing atomistic molecu-
191 lar dynamics software to carry out the coarse grained simulations, so that we do
192 not have to worry about writing efficient, parallelized simulation code. Moreover, at
193 this level the molecular shape is preserved which conceptually allows for a mapping
194 between the atomistic and coarse-grained representations of the system.

195 Motivated by Henderson’s theorem [29], which states that there exists a one-
196 to-one mapping between measured pair-correlation functions and the underlying
197 potential energy function in the case that the latter is a sum of pair-interactions,
198 strategies to construct coarse-grained forcefield are often based on constructing
199 effective pair-potentials from pair-correlation functions and potentials of mean force
200 obtained from atomistic simulations [30–34]. These approaches work well in cases
201 where 3-body and higher correlations are weak. The resulting effective potentials
202 that reproduce the target distributions are not unique however, which leaves room to
203 simultaneously match other target properties, for example experimental densities,
204 surface tensions, heats of vaporization, and so forth. Jain and co-workers showed
205 that such inclusion of thermodynamic target properties in the optimization proced-
206 ure has the added advantage of increased convergence [33]. A different method
207 to build coarse-grained potentials is by trying to match the effective forces on the
208 coarse-grained degrees of freedom within an atomistic simulation [35–37].

209 It is beyond the scope of the current multiscale topic to discuss the art of coarse-
210 graining in more depth. It is however important, and hopefully obviously so, to
211 ensure that the low-resolution and high-resolution models of the system, merged
212 in an hybrid multiscale method, represent the same thermodynamic state point.
213 The limited transferability of coarse-grained forcefields (as mentioned for $U(R)$ in
214 Eq. 2-1) as compared to atomistic forcefields, might therefore require one to re-
215 optimize the coarse-grained potentials for new simulation conditions (e.g. a different
216 temperature or pressure) to ensure the same chemical potential in the different low
217 and high resolution regions and avoid spurious density fluctuations.

218

219 **2.1.2. Mapping Between Different Representations, or the Reverse** 220 **Mapping Problem**

221

222 In order to implement an adaptive resolution algorithm, one must have a proced-
223 ure in place to dynamically change the representation of a molecule, either from
224 a finer to a coarser level or vice versa. The “forward” direction, namely where one
225 coarsens the representation of a molecule, is straightforward because one merely

226 throws away information; for example by replacing the atomic coordinates of a col-
227 lection of atoms with the coordinates of their center of mass (COM). On the other
228 hand, the “reverse” mapping, in which one resolves a molecule into finer detail,
229 is problematic because it requires the creation of information. Indeed, in a recent
230 review article on multiscale modeling, de Pablo and Curtin say that “a persistent
231 challenge remains that of reverse mapping, that is, of restoring some of the details
232 after they have been blurred away through an averaging procedure” [38].

233 Here we propose a solution to the reverse mapping problem. We associate a
234 frozen atomistic fragment with each coarse-grained site, and rigidly rotate these
235 fragments about their COM in accordance with an energy function designed to
236 maintain a low-energy atomistic-level representation of the system. In this manner
237 the coarse-grained system has, superimposed on it, a globally unfrustrated atomistic
238 configuration which is prepared for reverse mapping. Indeed, only a local relaxation
239 needs to be further performed on the degrees of freedom which were frozen (the
240 degrees of freedom within each fragment). This is accomplished, for instance, by
241 running a short atomistic simulation or by using a healing region as will be described
242 below. This method is efficient because the atomistic fragments are treated as rigid
243 bodies with no internal degrees of freedom. This means that there are only three
244 degrees of freedom per coarse-grain site to evolve, namely an element of the special
245 orthogonal Lie group $SO(3)$. The mathematical details of this treatment are given
246 below.

2.2. ADAPTIVE MULTISCALE MOLECULAR DYNAMICS

251 Due to the reverse mapping problem described above, constructing a meaningful
252 adaptive multiscale molecular dynamics algorithm is not trivial. First of all, the
253 number of degrees of freedom continuously changes during such a simulation as
254 low-resolution (coarse-grained) particles break up into their high-resolution compo-
255 nents (atoms) when they enter the high-resolution region and vice versa when they
256 leave. During these transitions the number of pair-interactions also changes, so that
257 both the total potential energy and the total kinetic energy show spurious fluctuations
258 with the changing numbers of particles in the low-resolution and high-resolution
259 regions. What thermodynamic ensemble would such an adaptive multiscale sim-
260 ulation sample? Secondly, and this is precisely the reverse mapping problem, a
261 coarse-grained particle entering the high-resolution region has to be replaced by its
262 fine-grained counterpart, which requires the generation of information: how can one
263 generate suitable equilibrium positions and velocities for these atoms in harmony
264 with the other atoms in the high-resolution region?

265 In this section, a possible solution is given to the above difficulties that come
266 with the construction of an algorithm for adaptive multiscale molecular dynam-
267 ics, or hybrid MD. The algorithm that we discuss is aimed to be simple, robust,
268 and have certain desired properties such as conservation of linear momentum and
269 conservation of total energy. This algorithm is built in two stages. Starting from
270 an atomistic (i.e. the high resolution) representation of the entire system and a

Multiscale Molecular Dynamics and the Reverse Mapping Problem

(forward) mapping of groups of atoms into coarse-grained beads, the atomistic pair-interactions that span atoms belonging to different coarse grained beads are replaced by coarse-grained pair-interactions in the coarse-grained region. The details of such a coarse-graining of interactions is presented in the following subsection where we discuss the treatment of the coupling between atoms in the atomistic and coarse-grained regions. The second stage then involves “freezing” of the atoms in the coarse-grained region inside their coarse-grained beads so that the dynamics of the atoms can be replaced by a coarse-grained dynamics. This is presented in the subsequent subsection. Next, we present an illustrative case study of a hybrid molecular dynamics (hybrid MD) simulation of liquid methane. We end with a brief discussion of variations of and alternatives to this multiscale algorithm that have recently appeared in the literature.

2.2.1. Stage 1: Coupling Atomistic and Coarse-Grained Regions

Let us consider a molecular system that we wish to separate into two regions, one of which contains the part that we want to treat in atomistic detail and the other which contains the rest to be treated at a lower, coarse-grained, resolution. This section describes the first of the two stages in which such a multiscale treatment can be constructed, leading to an adaptive algorithm that allows particles to move between the two regions and adapt their representation on the fly.

But first, let us recall that a conventional molecular dynamics treatment of the entire system at the atomistic level maintains the micro-canonical NVE ensemble as governed by the conservation laws of Hamilton’s equations. In this ensemble the number of atoms, N , the volume, V , and the total energy, E , remain constant. The total energy is taken to be the sum of the total kinetic energy, T , and the total potential energy, V , summed over all atoms, i :

$$E = T + V = \sum_i^N \frac{1}{2} m_i v_i^2 + V(r^N) \quad (2-2)$$

in which m_i and v_i are the mass and velocity of atom i respectively, and the potential depends on all the positions r_i .

Other ensembles can be generated, for example the canonical NVT ensemble or the isobaric-isothermal NPT ensemble by invoking a thermostat or a barostat plus a thermostat. Note that now the total energy is no longer constant. However, also in these ensembles a conserved quantity such as the total energy can be recovered by cleverly bookkeeping the energy flows between the molecular system and the external variables introduced to control the temperature and/or the pressure. The Nosé-Hoover (chain) thermostat [39, 40], the Parrinello-Rahman barostat [41], and the recently introduced stochastic velocity rescaling thermostat by Bussi [42] are all good examples of external control mechanisms that provide the desired ensemble and recover a conserved total energy even though the dynamics is not strictly Hamiltonian.

316 Conserved quantities play a very important role in molecular simulations.
 317 Monitoring the conservation of total energy yields the first and foremost assessment
 318 of the quality of the integration of the equations of motion during the simulation.
 319 Whether the computer code contains a bug, or the simulation time step is too large,
 320 or the initial system configuration contains overlaps between atoms, the problem is
 321 always first seen in the (non-) conservation of the total energy. Given this impor-
 322 tance, it seems like a good idea to focus on energy conservation while constructing
 323 a multiscale algorithm.

324 The construction of a multiscale algorithm begins from the atomistic representa-
 325 tion of the entire system and some mapping in which groups of atoms are lumped
 326 into coarse-grained beads, which can for example be positioned at the centers of
 327 mass of the atomic groups they represent. The first stage of coarse-graining the low-
 328 resolution part of the system consists of replacing all atomistic interactions that span
 329 different beads by coarse-grained interactions. The second stage consists of replac-
 330 ing the atoms in this region by their coarse-grained bead representation, which is
 331 dealt with in the next subsection. Specifically, for now, the atomistic bonded (and
 332 non-bonded) interactions between atoms i and j that belong to the same coarse-
 333 grained bead α are maintained, and only the interactions between atoms i and j
 334 that belong to different beads α and β are replaced by coarse-grained potentials. The
 335 energy function of such a system reads:

$$336 E = \sum_i^{n+m} \frac{1}{2} m_i v_i^2 + V^A(r^n) + V^{\text{CG}}(R^M) + V^{\text{mix}}(R^N, R^M) \quad (2-3)$$

$$337 V^A(r^n) = \sum_{i=1}^{n-1} \sum_{j>i}^n \Phi^A(r_i, r_j)$$

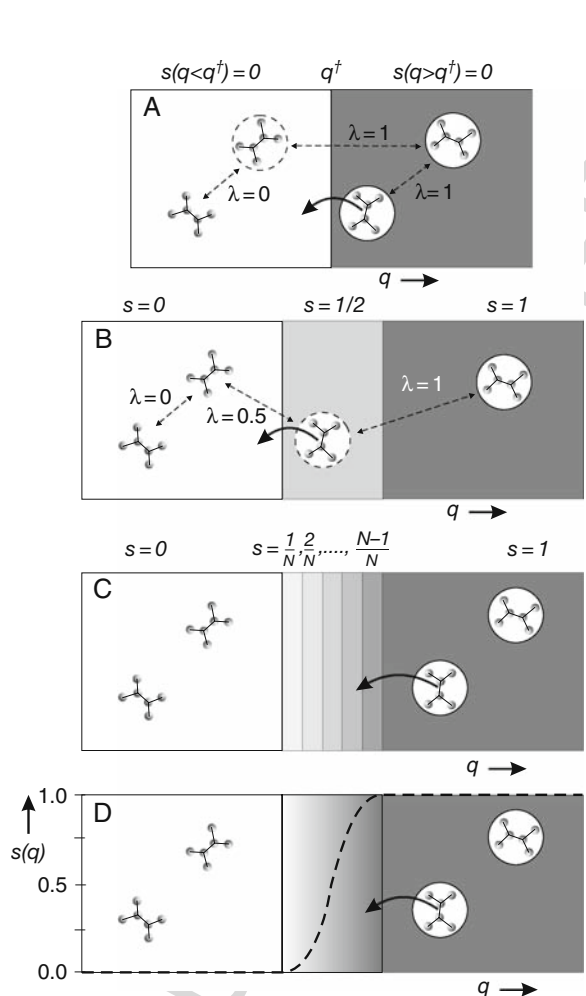
$$338 V^{\text{CG}}(R^M) = \sum_{\alpha=1}^{M-1} \sum_{\beta>\alpha}^M \Phi^{\text{CG}}(R_\alpha, R_\beta)$$

$$339 V^{\text{mix}}(R^N, R^M) = \sum_{\alpha=1}^N \sum_{\beta=1}^M \Phi^{\text{CG}}(R_\alpha, R_\beta)$$

340 so that the n atoms in the atomistic region interact with each other through
 341 an atomistic potential or forcefield $V^A(r^n)$ while the remaining m atoms in the
 342 coarse-grained region interact with each other through a coarse-grained potential
 343 or forcefield V^{CG} working on the M centers of mass, R , of the atomic fragments
 344 corresponding to beads. Here the potential terms are expressed as sums of pair-
 345 potentials, Φ . The effective force on such a center of mass is then distributed,
 346 mass weighted, over the atoms that belong to the coarse-grained bead. The cou-
 347 pling between atoms in different regions is governed by the last term $V^{\text{mix}}(R^N, R^M)$,
 348 which is also the coarse-grained potential working on the centers of mass and there-
 349 fore not different from V^{CG} , except that then α runs over the N beads in the atomistic
 350 region, instead of the $M - 1$ beads in the coarse-grained region.

Multiscale Molecular Dynamics and the Reverse Mapping Problem

361 The total energy in Eq. (2-3) is conserved, however this multiscale setup is not
 362 yet adaptive so that after some simulation time, due to diffusion, the atoms “feeling”
 363 fully atomistic interactions will start to mix with those atoms from the coarse-
 364 grained region that only feel their environment through coarse-grained interactions,
 365 which is not what we set out to do. To make the method adaptive and have the atoms
 366 switch their interaction when they cross user-defined regions, we introduce a spa-
 367 tial order parameter q to distinguish the two regions and we denote the boundary
 368 between the two regions by q^\dagger . With this order parameter all atoms can be attributed
 369 a label, s , which we shall call the (amount of) “coarse-grained character”, and which
 370 equals zero for atoms in the (atomistic) region of $q < q^\dagger$ and one for atoms in the
 371 (coarse-grained) region with $q > q^\dagger$, as shown in the upper panel (panel A) in
 372 Figure 2-1.



405 *Figure 2-1.* An intermediate “healing region” that smooths the coupling between the high-resolution and low-resolution regions is introduced in three simple steps

Next, the amount of coarse-grained character, s , is employed to make the potential function adaptive:

$$V^A + V^{\text{CG}} + V^{\text{mix}} = \sum_{\alpha\beta} \left(\lambda_{\alpha\beta} \Phi_{\alpha\beta}^{\text{CG}} + (1 - \lambda_{\alpha\beta}) \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A \right) + \sum_{\alpha} \sum_{i,j \in \alpha} \Phi_{ij}^A \quad (2-4)$$

$$\lambda_{\alpha\beta} = \max(s_{\alpha}, s_{\beta})$$

Here, the previous potential energy terms are replaced by the two terms on the right. The first term accounts for all interactions between different coarse-grained beads α and β , which are either taken as the coarse-grained pair-potential Φ^{CG} when $\lambda_{\alpha\beta} = 1$ or as the sum of atomistic pair-potentials Φ^A between atoms i belonging to bead α and atoms j belonging to bead β when $\lambda_{\alpha\beta} = 0$. The value of the scaling function, $\lambda_{\alpha\beta}$, being zero or one, is determined by the maximum coarse-grained character s between the two interacting beads α and β , as illustrated in Figure (2-1A). Using this scaling factor, the same multiscale behavior is obtained as before, namely, atoms in the atomistic region interact through the atomistic forcefield while they feel the particles in the coarse-grained region through the coarse-grained forcefield, and the atoms in the coarse-grained region “see” all other atoms, in both regions, as coarse-grained particles. The difference is now that this interaction automatically adapts when atoms move from one region to the other. The last term in Eq. (2-4) accounts for all bonded (and non-bonded) interactions within each bead. Note also that this equation is trivially generalized for interactions other than pair-interactions, such as bending (or dihedral) potentials, by taking for lambda the maximum s -value of the now three (or four) interacting atoms.

By making the multiscale approach adaptive, we have introduced two new problems: one, the total energy is no longer conserved, and, two, also the forces are discontinuous when particles cross the boundary at q^{\ddagger} . In other words, as long as the particles stay in their initial regions the energy is conserved and the dynamics is well-behaved, but every time a particle crosses q^{\ddagger} the total energy and the forces will display a jump. The jump in the energy when the atoms of bead α cross from the atomistic region to the coarse-grained region equals the difference between the coarse-grained potential and the atomistic potential due to all other atoms:

$$\Delta U_{\alpha}^{A/\text{CG}} = \sum_{\beta} \Delta \lambda_{\alpha\beta} \left(\tilde{\Phi}_{\alpha\beta}^{\text{CG}} - \sum_{\substack{i \in \alpha \\ j \in \beta}} \tilde{\Phi}_{ij}^A \right) \quad (2-5)$$

Multiscale Molecular Dynamics and the Reverse Mapping Problem

451 where $\Delta\lambda_{\alpha\beta}$ is the change of λ for each pair-interaction, which equals zero for all β
 452 particles in the coarse-grained region (as these λ remain unity), and for all other β
 453 equals either +1 if α moves from the atomistic region to the coarse-grained region,
 454 or -1 if α moves from the coarse-grained region to the atomistic region. The tilde
 455 indicates the value of the interaction, Φ , at the moment of boundary crossing. By
 456 simply adding this term, $\Delta U_{\alpha}^{A/CG}$, every time a particle crosses between regions,
 457 the total energy is recovered as a conserved quantity. Fixing the spurious jumps in
 458 the forces, however, will require more than just this simple bookkeeping. This is
 459 what we set out to do next.

460 Some improvement can be envisioned by introducing an intermediate “healing
 461 region” between the atomistic region and the coarse-grained region, in which par-
 462 ticles are attributed a mixed atomistic/coarse-grained character of $s = 0.5$. This is
 463 illustrated in Figure 2-1, panel B. Particles in this healing region interact with other
 464 “hybrid” particles in this region through a potential that is the sum of the atomistic
 465 interactions and the coarse-grained interactions, both scaled by $\lambda = 0.5$, according
 466 to Eq. (2-4). The same type of interaction is felt between these hybrid particles in the
 467 healing region and particle in the atomistic region (as $\max(s_1, s_2) = 0.5$), while the
 468 interaction between the hybrid particles and particles in the coarse-grained region is
 469 purely the coarse-grained Φ^{CG} (i.e. $\max(s_1, s_2) = 1$).

470 Similar to the previous dual-scale setup, the evolution of the atomistic dynam-
 471 ics can be performed on this, now three-region, system, where the potentials are
 472 switched accordingly every time a particle crosses the boundary between the atomis-
 473 tic region and the healing region or the boundary between the coarse-grained
 474 region and the healing region. In other words, the bookkeeping $\Delta U_{\alpha}^{A/CG}$ now sup-
 475 plements the energy function when a particle α switches between $s = 0$ and $s = 0.5$
 476 and when a particle switches between $s = 0.5$ and $s = 1$, to correct the jumps in
 477 the total energy. These corrections are smaller than with the previous single region
 478 boundary because $\Delta\lambda$ (in Eq. (2-4)) now only amounts to $+\frac{1}{2}$ or $-\frac{1}{2}$, and with that
 479 also the jumps in the forces are somewhat smaller.

480 Further improvement can thus be made by dividing the healing region into sev-
 481 eral, say $N-1$, intermediate sub-regions, as shown in panel C of Figure 2-1. Particles
 482 in each sub-region k could then be attributed a stepwise increasing coarse-grained
 483 character of $s = k/N$, starting from $k = 0$ in the fully atomistic region to $k = N$ in
 484 the fully coarse-grained region. Again, a particle in sub-region k interacts with all
 485 particles that find themselves in the same sub-region k or in regions smaller than k
 486 with a hybrid potential, given by Eq. (2-4), with scaling factor $\lambda = k/N$, whereas
 487 this particle interacts with all other particles in higher regions with a hybrid potential
 488 with a larger scaling factor that is determined by the other particle. For a fixed total
 489 healing region width with larger and larger numbers of intermediate sub-regions, N ,
 490 the intermediate regions become narrower in spatial extent and particles are more
 491 frequently found to cross one or another region boundary. A particle α crossing
 492 from region k to region $k + 1$ requires an update of the bookkeeping term $\Delta U_{\alpha}^{A/CG}$
 493 equal to:
 494
 495

$$\begin{aligned}
 \Delta U_{\alpha}^{A/CG} &= \sum_{\beta} \left(\frac{k+1}{N} \Phi_{\alpha\beta}^{CG} - \frac{k}{N} \Phi_{\alpha\beta}^{CG} - \frac{N-k-1}{N} \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A + \frac{N-k}{N} \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A \right) \\
 &= \sum_{\beta} \left(\frac{1}{N} \Phi_{\alpha\beta}^{CG} - \frac{1}{N} \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A \right)
 \end{aligned} \tag{2-6}$$

which is not different from the bookkeeping term defined by Eq. (2-4) with $\Delta\lambda = \pm 1/N$. In the case of very narrow sub-regions, a particle moving in any direction that is not strictly perpendicular to q will cross several sub-region boundaries, which requires summing over as many bookkeeping terms $\Delta U_{\alpha}^{A/CG}$.

Now, taking the limit of the number of sub-regions going to infinity,

$$\Delta U_{\alpha}^{A/CG} = \lim_{N \rightarrow \infty} \sum_{k_q}^{k_{q'}} \sum_{\beta} \Delta\lambda_{\alpha\beta} \left(\Phi_{\alpha\beta}^{CG} - \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A \right) \tag{2-7}$$

the sum over boundaries for a particle moving from sub-region k at q to sub-region k' at q' can be replaced by an integral:

$$\Delta U_{\alpha}^{A/CG} = \int_q^{q'} dq \sum_{\beta} \frac{d\lambda_{\alpha\beta}}{dq} \left(\Phi_{\alpha\beta}^{CG} - \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A \right). \tag{2-8}$$

Of course, the linear change of s and therefore of the scaling factor, λ , of $1/N$ per intermediate region is not a requirement, and other smoothly varying functions for s can be used, as long as it switches from zero to one over the healing region and is differentiable. During a molecular dynamics simulation, it is practical to perform the integration over the full spatial trajectory rather than its projection onto q , as both the interaction potentials, Φ , and the order parameter, q , are functions of the position, r .

$$\Delta U_{\alpha}^{A/CG} = \int dr \sum_{\beta} \frac{d\lambda_{\alpha\beta}}{dq} \frac{\partial q}{\partial r} \left(\Phi_{\alpha\beta}^{CG} - \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^A \right). \tag{2-9}$$

The resulting energy function of this adaptive multiscale method is obtained from Eq. (2-3), (2-4), and the total bookkeeping energy, $\Delta U^{A/CG}$, of all particles moving in the healing region. Integration of the equations of motion derived from this energy function, with a sufficiently small time step, will in principle maintain the total energy as a conserved quantity. In reference [7] an illustrative numerical calculation

Multiscale Molecular Dynamics and the Reverse Mapping Problem

is discussed of a simple model system containing a diatomic molecule which moves through a healing region. Despite the deliberately mismatched potential functions between the atomistic and coarse-grained representations, total energy conservation is recovered to remarkably high accuracy using the numerical integration of Eq. (2-9).

2.2.2. Equations of Motion

The Newtonian equations of motion of the adaptive multiscale molecular dynamics are constructed in the usual way

$$m_i \frac{d^2 r_i}{dt^2} = f_i = -\frac{\partial V}{\partial r_i} \quad (2-10)$$

with the force f on particle i derived from the potential,

$$V = \sum_{\alpha\beta} \left(\lambda_{\alpha\beta} \Phi_{\alpha\beta}^{\text{CG}} + (1 - \lambda_{\alpha\beta}) \sum_{\substack{i \in \alpha \\ j \in \beta}} \Phi_{ij}^{\text{A}} \right) + \sum_{\alpha} \sum_{i,j \in \alpha} \Phi_{ij}^{\text{A}} + \Delta U^{\text{A/CG}} \quad (2-11)$$

This is the adaptive multiscale, or hybrid MD, potential, which was derived in the previous section. Summarizing, the first term sums the scaled interactions between all pairs of coarse-grained particles α and β , with the scaling factor $\lambda_{\alpha\beta}$ a number between zero and one depending on the resolution (being atomistic, coarse-grained, or something in between) of the interacting particles. The second term accounts for all atomistic interactions within the coarse-grained beads. These interactions are not scaled by λ , as they are obviously not replaced by a coarse-grained interaction, and which would otherwise thus lead to disintegration of the molecules when they leave the atomistic region. The third and last term holds the bookkeeping energies (Eq. 2-9) of all particles that change their resolution in the healing region.

Derivation of the forces from the hybrid potential requires special attention to the first and last terms which contain the position dependent switching function. The force on atom i is

$$\begin{aligned} f_{i \in \alpha} &= -\frac{\partial V}{\partial r_i} \\ &= -\sum_{\beta} \left(\left\{ \lambda_{\alpha\beta} \frac{\partial \Phi_{\alpha\beta}^{\text{CG}}}{\partial r_i} + \frac{\partial \lambda_{\alpha\beta}}{\partial r_i} \Phi_{\alpha\beta}^{\text{CG}} \right\} \right. \\ &\quad \left. + \sum_{j \in \beta} \left\{ (1 - \lambda_{\alpha\beta}) \frac{\partial \Phi_{ij}^{\text{A}}}{\partial r_i} - \frac{\partial \lambda_{\alpha\beta}}{\partial r_i} \Phi_{ij}^{\text{A}} \right\} \right) \\ &\quad - \sum_{j \in \alpha} \frac{\partial \Phi_{ij}^{\text{A}}}{\partial r_i} - \frac{\partial \Delta U^{\text{A/CG}}}{\partial r_i} \end{aligned} \quad (2-12)$$

586 in which the terms in curly brackets are obtained by applying the chain rule to the
 587 scaled coarse-grained and atomistic interactions respectively, shown in Eq. (2-11)
 588 as the term in parenthesis. The first term in both of the pieces between the curly
 589 brackets we recognize as the usual interaction force, but now multiplied by the scal-
 590 ing factor λ and $(1 - \lambda)$ respectively. The second terms resulting from the chain rule
 591 contain the derivative of the scaling factor with respect to the particle position. This
 592 derivative is non-zero in the healing region where λ changes in the direction of the
 593 order parameter q (see also panel D in Figure 2-1). This seems strange as apparently
 594 these two terms will cause a force, and thus a flux of particles, between the atomistic
 595 and coarse-grained regions, unless $\Phi_{\alpha\beta}^{\text{CG}} = \sum_{ij} \Phi_{ij}^{\text{A}}$.

596 However, the last term, the derivative of $\Delta U^{\text{A/CG}}$ is just the integrand of
 597 Eq. (2-9)

$$598 \frac{\partial \Delta U^{\text{A/CG}}}{\partial r_i} = \sum_{\beta} \left(\frac{\partial \lambda_{\alpha\beta}}{\partial r_i} \Phi_{\alpha\beta}^{\text{CG}} - \sum_{j \in \beta} \frac{\partial \lambda_{\alpha\beta}}{\partial r_i} \Phi_{ij}^{\text{A}} \right) \quad (2-13)$$

603 which cancels exactly these two spurious terms leaving only the scaled interactions
 604 plus the intra-bead interaction in the force expression:

$$605 f_{i \in \alpha} = - \sum_{\beta} \left(\lambda_{\alpha\beta} \frac{\partial \Phi_{\alpha\beta}^{\text{CG}}}{\partial r_i} + \sum_{j \in \beta} (1 - \lambda_{\alpha\beta}) \frac{\partial \Phi_{ij}^{\text{A}}}{\partial r_i} \right) - \sum_{j \in \alpha} \frac{\partial \Phi_{ij}^{\text{A}}}{\partial r_i} \quad (2-14)$$

606 This equation contains the usual symmetry with respect to interacting particles i
 607 and j which ensures obedience to Newton's third law ($f_i = -f_j$) and conservation of
 608 momentum in the system.

614 2.2.3. Stage 2: Freezing the Intra-Bead Motions

615 In the previous two sections an adaptive multiscale dynamics approach was con-
 616 structed by coupling an atomistic region, in which atoms interact through an
 617 atomistic forcefield, with a coarse-grained region in which atoms interact through
 618 a coarse-grained forcefield. Technically, however, both regions still maintain atom-
 619 istic dynamics rather than a coarse-grained dynamics in the coarse-grained region.
 620 That is, the integrator propagates the atomic positions and velocities also in the
 621 coarse-grained region instead of propagating positions and velocities of the coarse-
 622 grained particles. Here, in the second stage of the multiscale method development,
 623 the atoms in the low-resolution region are replaced by coarse-grained beads.

624 Basically, the atoms grouped into a coarse-grained bead can be "frozen" with
 625 respect to their center of mass and replaced by the coarse-grained particle as soon
 626 as they enter the coarse-grained region from the healing region, since, from there
 627 on, the motion of the center of mass of each atomic fragment is solely governed by
 628 the coarse-grained interactions making the "internal" atomic motions irrelevant to
 629 the molecular dynamics. Such one-step-freezing is exactly what we will do here,
 630

Multiscale Molecular Dynamics and the Reverse Mapping Problem

631 although alternative schemes are possible, involving for example a gradual freezing
632 of the atomic degrees of freedom in the healing region (see section for more details).

633 The advantages of replacing the evolution of the atomic degrees of freedom by
634 coarse-grained molecular dynamics are mainly economic; computation of atomic
635 intra-bead interactions is avoided, a larger time step is allowed without violat-
636 ing energy-conservation, and memory storage can be saved in the coarse-grained
637 region, which is typically the larger region. The only disadvantage of discarding
638 the atomistic details is that of the reverse mapping problem discussed earlier: once
639 the atomic inter-bead interactions are replaced by the coarse-grained forcefield the
640 atomic fragments are free to rotate around their center of mass, randomizing the
641 atomistic details with respect to their environment outside the bead. This is why
642 we in fact control this rotation as discussed in section , so that we can recover the
643 atomic details to some approximation.

644 The fact that, in the low-resolution region, the atoms no longer feel their envi-
645 ronment and adopt random orientations (ignoring for the moment the possibility of
646 SO(3) rotational dynamics) means that such groups, upon moving into the healing
647 region and toward the atomistic region, practically always do so starting relatively
648 high up on the atomic potential energy surface. That is, while the atomic interactions
649 are gradually switched on while moving toward the atomistic region, the atomic
650 degrees of freedom (have to) re-equilibrate with respect to their atomistic environ-
651 ment. This equilibration process of transforming from high potential energy random
652 orientations to equilibrium energy structures generates thermal motion or heat. Note
653 that the reverse process of switching off the atomistic interactions and allowing
654 atomic fragments to take random orientations as molecules move toward the coarse-
655 grained region does not require the absorption of heat. This asymmetry in heat
656 transport with respect to movement toward the high-resolution region versus move-
657 ment in the other direction means that heat is produced continuously in adaptive
658 multiscale molecular dynamics. This heat has to be removed by a thermostat.

659 When particles cross the boundary between the coarse-grained region and the
660 healing region, the atoms are replaced by a single coarse-grained particle or vice
661 versa depending on the crossing direction. In the coarse-grained region, the atomic
662 positions and velocities can simply be stored relative to their center of mass and
663 thus be recovered when the particle leaves the coarse-grained region. Alternatively,
664 atomic positions and velocities can be re-generated, for example by inserting a
665 relaxed structure with random (Boltzmann) velocities or by copying positions and
666 velocities from a molecule in the atomistic region. The instantaneous switching
667 between atoms and coarse-grained particles at this boundary introduces jumps in
668 the total energy that require two extra bookkeeping terms, namely

$$669 \Delta U^{\text{intraCG}} = \sum_{\alpha} \Theta(s_{\alpha} - 1) \sum_{i,j \in \alpha} \tilde{\Phi}_{ij}^A \quad (2-15)$$

$$670 \Delta T^{\text{A/CG}} = \sum_{\alpha} \Theta(s_{\alpha} - 1) \frac{1}{2} \left(m_{\alpha} \tilde{v}_{\alpha}^2 - \sum_{i \in \alpha} m_i \tilde{v}_i^2 \right) \quad (2-16)$$

676 The first extra bookkeeping term accounts for the atomistic intra-coarse-grained
 677 bead interactions which are no longer computed when the atoms become “frozen”.
 678 Here, the tilde indicates again the values of Φ and ν at the boundary crossing
 679 moment and the Heaviside step function, Θ , equals one for particles in the coarse-
 680 grained region ($s = 1$) and zero otherwise. The second equation accounts for the
 681 change in kinetic energy as the number of degrees of freedom adapts. The sum of
 682 the two terms can be seen as the instantaneous internal energy of the atomic degrees
 683 of freedom inside the coarse-grained bead, which is integrated out upon coarse-
 684 graining and thus becomes a constant in the coarse-grained region, whereupon it
 685 will be released again when switching back to the atomistic representation.

686 The definition of the regions (high-resolution, healing, and low-resolution) is
 687 arbitrary and can for example be chosen to be a spherical atomistic region centered
 688 on a particle of particular interest, so that the high-resolution part of the simulation
 689 follows this particle. The healing region should then also be a spherical region with
 690 the same center but with a larger radius, leaving everything outside this sphere as the
 691 low-resolution region. Alternatively, the regions can be fixed in space. In either case,
 692 after every (smallest) time step the amount of coarse-grained character, s , needs to
 693 be computed by calculating the distance between each particle and the center of
 694 the spherical regions. The coarse-grained character is used for the scaling factor in
 695 the force calculations involving all particles in the healing region. A computational
 696 saving can be made by computing s only for a list of particles that find themselves
 697 in this healing region or just outside of it (skin-regions). By taking the width of the
 698 skin-regions on either side of the healing region equal to the skin length applied in
 699 the usual neighbor list for the non-bonded interactions, the small list of particles for
 700 which s is computed needs to be rebuild only as often as the neighbor list is updated.

701 Summarizing, we have constructed an adaptive multcale molecular dynamics
 702 by first introducing an intermediate healing region and subsequently replacing
 703 the groups of atoms in the low-resolution region by coarse-grained beads. The
 704 total energy is recovered as a conserved quantity by adding the proper auxiliary
 705 bookkeeping terms to the energy function. Next, we will examine the behavior of
 706 this hybrid MD approach with an illustrative example.

707

708

2.2.4. Case Study 1: Liquid Methane

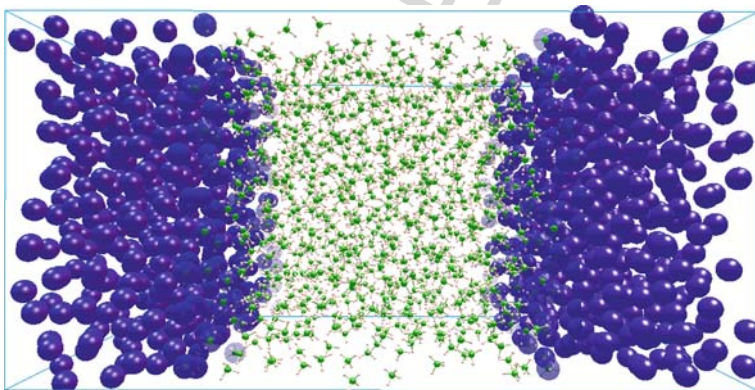
709

710 As an illustrative application of the adaptive multiscale dynamics method, we have
 711 performed a simulation of liquid methane at $T = 111.5$ K and atmospheric pressure.
 712 In the high resolution region, methane is treated as a fully flexible atomistic CH_4
 713 molecule (using the CHARMM forcefield [43]) and in the low resolution region
 714 each methane molecule is modeled as a single bead using Jorgensen’s united atom
 715 model [44]. A similar simulation of 8000 methane molecules has been published
 716 in reference [7], in which case the atomistic region was defined by a sphere with
 717 a radius of $R_A = 8 \text{ \AA}$ fixed in space. Different surrounding healing regions were
 718 tried with widths varying between $R_{HR} = 1$ and 4 \AA , and, not surprisingly, the
 719 largest, 4 \AA wide, healing region resulted in the best performance as seen from the
 720 conservation of energy.

Multiscale Molecular Dynamics and the Reverse Mapping Problem

721 Here, we take instead a rectangular box with edges $L = 38.0, 38.0, 75.5$ Å, con-
722 taining 1900 methane molecules subdivided into rectangular regions of different
723 resolution. That is, the atomistic region is a slab, flanked by healing regions on both
724 sides with a total thickness of $d_A + 2 \times d_{HR} = 44$ Å, which leaves the rest as the
725 system as a coarse grained slab of width $d_{CGR} = 31.5$ Å. See Figure 2-2 for an
726 illustration of the system. We will compare six hybrid MD simulations in which we
727 again vary the width of the healing regions from $d_{HR} = 1$ to 6 Å (thus leaving an
728 atomistic region of varying width between $d_A = 42$ and 32 Å). The average number
729 of methane molecules in the coarse-grained region was close to 790 with a standard
730 deviation of about 17. The number of molecules in the atomistic region varied from
731 1063 in the $R_{HR} = 1$ Å simulation to 812 in the $R_{HR} = 6$ Å simulation, leaving
732 50–294 molecules respectively in the healing regions (see also top-right panel in
733 Figure 2-3).

734 The interesting observables that illustrate the behavior of the hybrid MD method
735 are plotted in the other three panels of Figure 2-3. Starting at the top-left panel, we
736 see the total energy (shifted for comparison) is very well conserved for the $R_{HR} = 6$
737 Å healing region, and showing an unstable drift in the hybrid MD simulations with
738 healing region smaller than $R_{HR} = 3$ Å. The graph at the bottom left shows the
739 $\Delta U^{A/CG}$ bookkeeping term (Eq. 2-9), which in the simulations is computed on the
740 fly by multiplying the integrand (Eq. 2-13) by the displacement dq of the interacting
741 particle in the healing region. Molecules moving from the coarse-grained region to
742 the atomistic region fall quickly down from high potential energy configurations as
743 they equilibrate into their atomistic environment, whereas molecules moving in the
744 other direction are not *pushed up* to such high potential energy configurations. This
745 asymmetry with respect to the direction that particles move (i.e. positive or negative
746
747
748



749
750
751
752
753
754
755
756
757
758
759
760
761 *Figure 2-2.* Hybrid MD snapshot of the periodic unit cell of 1900 methane molecules. In the center is
762 the atomistic region flanked by the two halves of the coarse-grained slab, in which each CH_4 molecule
763 is treated as a single blob (pictured in *blue*). These two regions are coupled through 6 Å wide healing
764 regions in which the molecules smoothly change their resolution, here indicated by the varying
765 transparency of the molecules

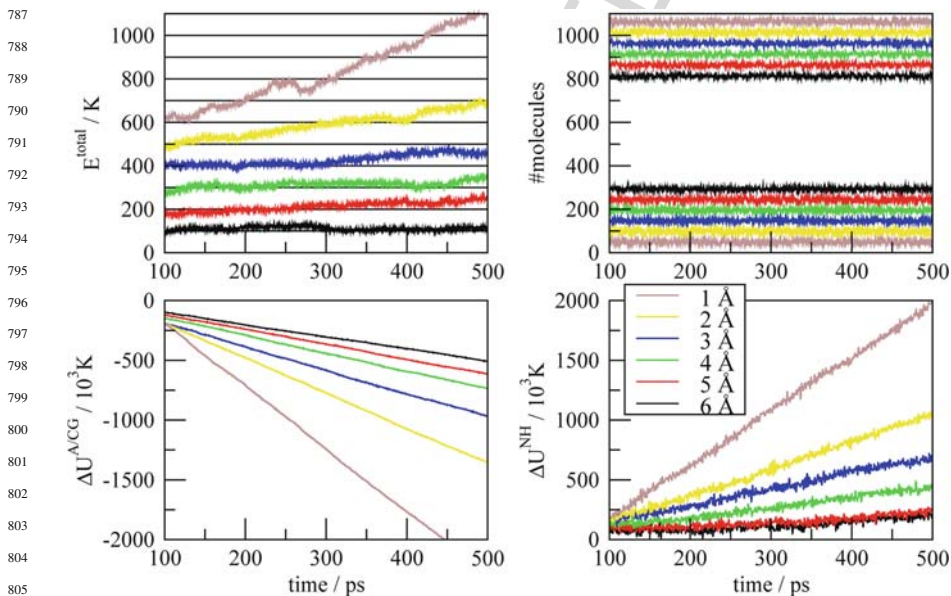
766 dq in Eq. 2-9) means that on average the fluctuations in $\Delta U^{A/CG}$ do not cancel, as
 767 reflected by the negative slopes in the plot.

768 The atomistic equilibration process in the healing region, driven by a gradual
 769 switching on of the atomistic interactions, forces the atoms to move and reorient.
 770 This increase in kinetic energy is counteracted by a Nosé-Hoover thermostat coupled
 771 to each particle. The bottom-right panel in Figure 2-3 shows the flow of energy from
 772 the system to the thermostat, which is larger for small healing region widths.

773 This case study of a hybrid MD simulation of liquid methane illustrates that the
 774 adaptive multiscale algorithm is a robust and very promising method. It also shows
 775 the importance of recovering the total energy as a conserved quantity in assessing the
 776 stability and accuracy of the simulation. Without this total energy observable, one
 777 is left with guessing, based on secondary information such as density fluctuations
 778 or other measured properties, whether the choices made for the healing region size
 779 and the time step were adequate.

781 2.2.5. Other Adaptive Multiscale Implementations

782 At this stage, we briefly outline two other approaches to (particle-based) adap-
 783 tive multiscale dynamics methods, paying special attention to the differences and
 784 similarities to the hybrid MD method constructed in the previous sections.
 785



806 *Figure 2-3.* Hybrid MD simulation data for liquid methane with varying healing region widths. *Top left:*
 807 the conservation of the total energy is excellent for the 6 Å healing region and becomes unacceptable
 808 for healing regions smaller than 3 Å. *Top right:* Continuously fluctuating numbers of molecules in the
 809 atomistic region (*upper part* of the graph) and in the healing region (*lower part*). *Bottom left:* bookkeep-
 810 ing term $\Delta U^{A/CG}$ from Eq. 2-9. *Bottom right:* continuous energy flow into the Nosé-Hoover thermostats

Multiscale Molecular Dynamics and the Reverse Mapping Problem

811 Praprotnik et al. [8] were the first to present an adaptive resolution dynamics
 812 scheme, only preceded by the similar adaptive Monte-Carlo approach of Abrams
 813 [45]. Instead of starting from an energy function with scaled potentials (as we did
 814 in Eq. 2-4), Praprotnik uses force scaling:

$$815 \quad f_{\alpha\beta} = w(R_\alpha)w(R_\beta) \sum_{i\alpha,j\beta} f_{i\alpha j\beta}^A + [1 - w(R_\alpha)w(R_\beta)]f_{\alpha\beta}^{CG} \quad (2-17)$$

816 Here, w are the scaling functions of the interacting beads α and β . This scheme
 817 also obeys Newton's third law and is constructed not to cause any flux of par-
 818 ticles over the intermediate healing region. Unfortunately, the energy function is
 819 not known in this approach. Another difference is the use of fractional degrees
 820 of freedom. The temperature is then calculated using the fractional analog of the
 821 equipartition theorem
 822

$$823 \quad \langle K_\alpha \rangle = \frac{w(R_\alpha)k_B T}{2} \quad (2-18)$$

824 where $\langle K_\alpha \rangle$ is the average kinetic energy per fractional degree of freedom
 825 [46, 47]. In this approach, all interactions, including the atomistic intra-bead inter-
 826 actions are scaled, and simultaneously the kinetic energy of the atomic degrees of
 827 freedom is scaled to zero when a particle switches from atomistic to coarse-grained
 828 resolution. Control of the kinetic energy is established through a position depen-
 829 dent dissipative particle dynamics thermostat coupled to each (fractional) degree of
 830 freedom.

831 The other approach worth mentioning is to derive a mixed-resolution Hamil-
 832 tonian, starting from a linear combination of all possible Lagrangians that can be
 833 constructed by treating all particles in the atomistic region plus a subset of those in
 834 the healing region at the high-resolution level of theory, and treating all other par-
 835 ticles at the low resolution [9]. Also in this case, the internal, high-resolution, degrees
 836 of freedom gradually "freeze" when particles leave the atomistic region, through
 837 a coordinate dependent kinetic energy. Although this approach is somewhat more
 838 involved, requiring the implementation of a special generalized integrator and the
 839 definition of a special mixed-resolution potential energy function, one advantage is
 840 that it is not limited to pair-potential forcefields but can also be used in combination
 841 with many-body forcefields, such as is common in a quantum mechanical treatment
 842 [48]. Also, the total momentum and total energy are strictly conserved.
 843

844 2.3. REVERSE MAPPING THROUGH RIGID BODY ROTATION

845 From the viewpoint of the reverse mapping problem, the example of liquid methane
 846 shown in the previous section represents the simplest possible case. The atomistic
 847 methane molecule, CH_4 , is symmetrical: it has four hydrogen atoms in a tetrahedral
 848 arrangement around a central carbon atom, and hence to a first approximation is
 849 well-described by the united atom sphere which is used to represent it in the coarse
 850 grained region. When such a united atom sphere enters the healing region, one may

856 conceivably insert the missing hydrogen atoms in a randomly oriented tetrahedron
857 without paying too steep a price in potential energy. The tetrahedron will need to
858 rotate to avoid steric clashes with nearby methane molecules, and this motion is
859 the source of the energy flow to the thermostat seen in the previous section. These
860 steric clashes are non-bonded in nature, and non-bonded forces are the softest forces
861 which exist in the forcefield. Soft forces are easily “healed” in the healing region.
862 Once one moves from methane to more complex molecules, the reverse mapping
863 problem can become serious, and one needs a more sophisticated strategy to solve it.

864 Recall that our solution to the reverse mapping problem is to associate a frozen
865 atomistic fragment with each coarse-grain site, and to rigidly rotate these fragments
866 about their COM in accordance with an energy function designed to maintain a
867 low-energy atomistic-level representation of the system. The atomistic fragments
868 are treated as rigid bodies with no internal degrees of freedom. This means that
869 there are only three degrees of freedom per coarse-grain site to evolve, namely an
870 element of the special orthogonal Lie group $SO(3)$. In the remainder of this section
871 we present several strategies for implementing this $SO(3)$ rotational motion. First
872 we present an energy minimization approach using conjugate gradient optimization.
873 Next, we describe a dynamics approach using the direct analog of the velocity Verlet
874 algorithm for the $SO(3)$ setting. We then discuss the coupling between the $SO(3)$
875 dynamics and the coarse-grained molecular dynamics, and finish this section with a
876 case study of a polyethylene chain.

878 2.3.1. Rigid Body Rotational Optimization

880 Here we present an algorithm that uses $SO(3)$ optimization to align molecular frag-
881 ments corresponding to coarse-grained sites. The output from this algorithm consists
882 of rigid fragments centered at the coarse grained sites, rotated to minimize an
883 energy function consisting of both intra- and inter-molecular terms. The approach
884 is based on an algorithm due to Taylor and Kriegman [49] in which a sequence
885 of local parameterizations of the manifold $SO(3)$ is used, rather than relying on a
886 single global parameterization such as the Euler angles. The problems caused by
887 singularities in a global parameterization are thus avoided.

888 One can object to an energy-minimized structure on the grounds that it is not
889 compatible with the ensembles typically used in molecular dynamics simulations
890 (e.g. NVT or NPT). In this sense, one can ask the question “In what sense is this
891 configuration a representation of the underlying molecular system?” Rather than
892 attempting to provide a mapping algorithm that generates an equilibrated atomistic
893 configuration, in this section we take a more pragmatic approach. The mapping
894 algorithm described here quickly generates a globally stable atomistic configuration
895 that further requires very localized relaxation and equilibration.

896 The algorithm minimizes a real-valued objective function $E:SO(3) \rightarrow \mathbb{R}$ defined
897 on the set of rotation matrices

$$898 R \in SO(3) \equiv \left\{ \in \mathbb{R}^{3 \times 3} : R^t R = I, \det(R) = 1 \right\}. \quad (2-19)$$

899

900

Multiscale Molecular Dynamics and the Reverse Mapping Problem

At every point R_0 on the manifold $SO(3)$ we construct a continuous, differentiable mapping between a neighborhood of R_0 on the manifold and an open set in \mathbb{R}^3 ,

$$R(\omega) = R_0 \exp J(\omega), \omega \in \mathbb{R}^3, |\omega| < \pi \quad (2-20)$$

where the skew symmetric operator $J: \mathbb{R}^3 \rightarrow SO(3)$ is defined as

$$J(\omega) = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}. \quad (2-21)$$

$R(\omega)$ can be computed using the Rodrigues formula, although we will not need to do so. The objective (energy) function can be expanded to quadratic order about R_0 as

$$E(R(\omega)) = E(R_0) + g^t \omega + \omega^t H \omega \quad (2-22)$$

where g and H are the gradient and the Hessian of the function, respectively, evaluated at the point $\omega = 0$ which corresponds to the rotation matrix R_0 . The conjugate gradient incremental step is

$$\omega_s = -H^{-1} g. \quad (2-23)$$

This incremental step determines the new rotation matrix as follows:

$$R = R_0 \exp J(\omega_s). \quad (2-24)$$

The incremental step must lie within the range of the local parameterization, namely $|\omega_s| < \pi$. The updating step can be made computationally efficient by representing the rotations by unit quaternions. The relationship between $SO(3)$ and the group of unit quaternions $Sp(1)$ is

$$q = (\cos \theta, \hat{\omega} \sin \theta), \theta = |\omega|/2. \quad (2-25)$$

The incremental step corresponds to the quaternion

$$q_s = \left(\cos \frac{\theta}{2}, \omega \frac{\sin(\theta/2)}{\theta} \right) \text{ where } \theta = |\omega|. \quad (2-26)$$

With the rotation R_0 expressed as the unit quaternion q_0 , the product of the two rotations, which gives the new rotation matrix, is given by the quaternion multiplication $q_0 q_s$. It has been shown that this algorithm exhibits quadratic convergence provided that the starting point is sufficiently close to a minimum. According Eq. (2-22), we are supposed to evaluate the Hessian as well as the gradient to compute the update step. However, by employing the Fletcher-Reeves-Polak-Ribiere version of the conjugate gradient algorithm, only the gradient is needed [50]).

To apply this algorithm to molecular systems, two things are needed. Firstly, an objective function must be chosen which imbues the algorithm with chemical meaning. This function should provide a measure of the potential energy of the molecular configuration associated with a given rotation matrix. Secondly, the algorithm must be extended to many coupled SO(3) optimizations so that the molecular system is simultaneously and concertedly optimized over all the coarse-grain centers. This multi-body extension is in fact trivial and does not incur any additional computational cost aside from the necessary linear scaling with the number of centers. The nature of the multi-body aspect of the algorithm will become clear in what follows.

Let us now address the choice of an energy function. Only interactions between atoms belonging to different coarse-grain units need be considered. This is because the intra-unit degrees of freedom are frozen. The SO(3) algorithm is designed to find the optimal rotational orientation of each of these fragments, where the center of mass of each fragment is constrained to lie at the location of the coarse-grained site representing it, and where no internal relaxation of the intra-fragment degrees of freedom is allowed. Ideally, we would like to take the functional form and the parameters of all of the terms contributing to the energy function from an underlying atomistic force field. There is no need to invent new potential energy terms when we have well-parameterized ones at our disposal. The most important contribution to the energy function is a bonded term arising from the “dangling” bonds in the molecular fragments which would connect the fragments to one another in an atomistic representation of the system. This function is harmonic in the interatom distance and is expressed as (see Figure 2-4)

$$E(R_1, R_2) = k/2 (|r + R_2v - R_1u| - d_0)^2 \quad (2-27)$$

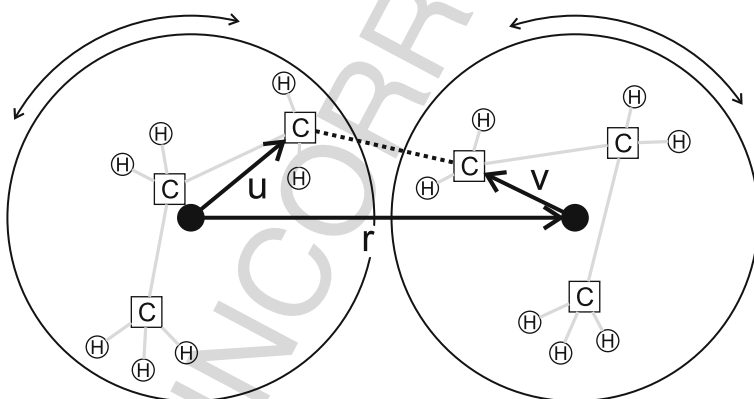


Figure 2-4. Schematic of the SO(3) optimization algorithm for a hexane molecule. Atomistic fragments are rotated about the centers of mass of the coarse-grain units representing them to align the unconnected atomistic bond between fragments (shown in *dashed line*)

Multiscale Molecular Dynamics and the Reverse Mapping Problem

where R_1 and R_2 are the rotation matrices corresponding to coarse grain units 1 and 2, located at positions COM1 and COM2, respectively. The vector from COM1 to COM2 is denoted r . u represents the vector from COM1 to the atom in coarse-grain unit 1 involved in the bond. v represents the vector from COM2 to the atom in coarse-grain unit 2 at the other end of the bond. The bond has an equilibrium distance of d_0 and a force constant of k . In order to perform optimization, the gradient must be evaluated. There are six gradient terms associated with this energy function, three for rotation matrix R_1 and three for R_2 . For the ω_x component of R_1 , the gradient is

$$\frac{\partial}{\partial \omega_{x1}} E = -k \frac{(|r + R_2 v - R_1 u| - d_0)}{|r + R_2 v - R_1 u|} (r + R_2 v - R_1 u) \cdot \frac{\partial}{\partial \omega_{x1}} R_1 u \quad (2-28)$$

with

$$\frac{\partial}{\partial \omega_{x1}} R_1 u = R_0^1 J(\hat{x}) u \quad (2-29)$$

where R_0^1 denotes the initial rotation matrix for coarse-grained center 1. This last result is computationally important because it means that the J operator only ever acts on three elements, \hat{x} , \hat{y} , and \hat{z} , where,

$$J(\hat{x}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, J(\hat{y}) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, J(\hat{z}) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (2-30)$$

The multi-body nature of the minimization procedure proceeds as follows: the potential energy function is written as a sum over all the coarse-grain sites, with separate terms arising from bonds, bends, torsions, one-fours, and Lennard-Jones interactions. The gradient of this function is evaluated with respect to each degree of freedom, namely the three numbers ω_x , ω_y , and ω_z for each coarse-grain site. This gradient is used to decide upon a global incremental update step in which all the rotation matrices are simultaneously changed. For further details on how to implement bend, torsional, and non-bonded energy terms, we refer the interested reader to our previous work [51].

2.3.2. Rigid Body Rotational Dynamics

In the interest of energy conservation, we now turn to rigid body rotational dynamics algorithms, namely how to adapt the velocity Verlet algorithm for the setting of the SO(3) Lie group. This is a challenging problem because the dynamics occurs on a curved manifold, not a vector space. There are no known algorithms which possess all the desired properties expected of such an integrator, namely that it be symplectic, time-reversible, and energy and momentum conserving. We have chosen to implement the best-performing currently known explicit algorithm for rigid body dynamics, which is due to Krysl and Endres [52]. This algorithm is the direct analog

of the velocity Verlet algorithm for the rotation dynamics setting. The fundamental law of motion for rotational dynamics is $t = \dot{L}$ where t is the torque, $L = I\omega$ is the angular momentum, I is the moment of inertia tensor, and ω is the angular velocity. The initialization and subsequent dynamics components of the algorithm are as follows.

We initialize the angular velocity from the coordinates and (linear) velocities of the atoms via

$$\omega_0 = I_0^{-1}L_0 = I_0^{-1}\left(\sum_i r_i \times m_i v_i\right), \quad (2-31)$$

where I is computed relative to the COM, r_i is the vector from the COM to atom i ; m_i is the mass of atom i , v_i is the (linear) velocity of particle i , and \times is the cross product. The angular acceleration is initialized from

$$\alpha_0 = I_0^{-1}(t_0 - \omega_0 \times I_0 \omega_0), \quad (2-32)$$

where t is the torque.

Following initialization, the dynamics occurs in a 4-step algorithm as follows, where the time step is τ .

Step 1: $\omega'_n = \omega_{n-1} + \frac{\tau}{2}\alpha_{n-1}$

Step 2: $R_n = \exp[\tau J(\omega'_n)]$

Step 3: Solve $I_n \alpha_n - t_n + (\omega'_n + \frac{\tau}{2}\alpha_n) \times I_n (\omega'_n + \frac{\tau}{2}\alpha_n) = 0$ for α_n

Step 4: $\omega_n = \omega'_n + \frac{\tau}{2}\alpha_n$

Steps 1 and 4 are the velocity half-steps. In Step 2 the rotation matrix R_n is used to update the atomic positions relative to their COM. Hence, unlike the original algorithm, we do not apply incremental updates to the rotation matrix, but instead directly rotate the atoms at each time step. We remind the reader that $J(\omega)$ is defined by Eq. (2-21). In between Steps 2 and 3, the torque and the moment of inertia tensor are updated. In Step 3 we need to solve a non-linear vector equation for the angular acceleration α_n using Newton's method since it is coupled with the angular velocity due to the velocity update. The Jacobian (the matrix of first partial derivatives) required for Newton's method is straightforward to evaluate [52].

The multi-body nature of the algorithm, as was the case for the optimization algorithm presented above, is trivial: the net torque on each coarse-grained center is computed due to the effect of all other fragments one wishes to consider.

2.3.3. Coupling Between the Rotational Dynamics and Coarse-Grained Molecular Dynamics

To allow for reverse mapping on the fly during a coarse-grained molecular dynamics simulation, we wish to couple the velocity Verlet algorithm introduced above for the rigid body rotational dynamics to the dynamics of the coarse-grained centers of mass. That is, we wish to perform the rotational dynamics of the atoms within

Multiscale Molecular Dynamics and the Reverse Mapping Problem

1081 the coarse-grained beads simultaneously with the (Cartesian) molecular dynamics
1082 of the beads in a fashion reminiscent of Car-Parrinello MD. In the latter method,
1083 the classical dynamics of the nuclei are in principle coupled to a second, artificial,
1084 dynamics of electronic wave function degrees of freedom, although in that case
1085 an adiabatic separation exists due to the large difference in particle masses, and
1086 therefore in temperatures, between the two sub-systems. Here, such a separation
1087 does not exist.

1088 The orientation of the rigid body and its rotational motion are governed
1089 through atomistic interactions that span different beads as illustrated by the dashed
1090 C–C bond in Figure 2-4. These atomistic interactions are affected if we allow
1091 the beads to move with respect to each other. Imagine for example that the
1092 two beads in Figure 2-4 are moving away from each other, then clearly also the
1093 (dashed) C–C bond elongates, leading to an increased torque on the atomistic body
1094 and thus a speedup of the rotational motion. That means in practise that the rota-
1095 tional dynamics continuously heats up if we allow the beads to move, unless a
1096 corresponding back-coupling of the rigid body rotation is added to the dynamics
1097 of the beads or, alternatively, a friction is added to the rotational dynamics to avoid
1098 heating up. The latter damped dynamics is most easily implemented and results
1099 in an alternative on-the-fly rotational optimization scheme (see also Section); this
1100 solution is an example of a general technique known as simulated annealing. Here
1101 instead, we will discuss the (back-) coupling between the (undamped) dynamics
1102 subsystems, as it will play a role in the hybrid multiscale method (see Section 2.1).

1103 Adding the correct back-coupling of the rotational dynamics to the dynamics of
1104 the beads entails adding the atomistic interactions that govern the rotational motions
1105 to the coarse-grained dynamics. However, we do not wish to disrupt the coarse-
1106 grained molecular dynamics more than necessary, not in the least because adding
1107 stiff atomistic interactions to the coarse-grained dynamics would require a smaller
1108 time step for its evolution. We therefore distinguish between translational and rota-
1109 tional motion of the beads with respect to each other, where we define translation as
1110 the motion that alters the distance between two beads and rotation as the motion in
1111 which the bead-bead distance remains constant.

1112 For the translational motion of beads (imagine the coarse-grained stretch vibra-
1113 tion between the two beads in Figure 2-4), instead of adding the back-coupling
1114 of the atomistic interaction to the coarse-grained dynamics, we remove the cou-
1115 pling of the coarse-grained translational motion to the rotational dynamics. This is
1116 done by re-normalizing the atomistic configurations to the equilibrium bead-bead
1117 distance, instead of using the actual bead-bead distance. That is, the instantaneous
1118 bead-bead distance r in Eq. (2-27) for the rotational potential energy is replaced by
1119 the equilibrium bond distance r_0 , as follows

$$1121 \quad E(R_1, R_2) = k/2 (|r_0 + R_2 v - R_1 u| - d_0)^2 \quad (2-33)$$

1122
1123
1124 and likewise in the force expression of Eq. (2-28). This way, the rotation dynamics
1125 is computed as if the beads are always placed at their equilibrium distance, thus

1126 removing any effect of translational motion on the rotational dynamics and the need
1127 to add back-coupling.

1128 For the rotational motion of the beads, we cannot remove the coupling because it
1129 would defeat the purpose of keeping the atomistic configuration while the (coarse-
1130 grained) molecule rotates. That is, when the coarse-grained molecule rotates, we
1131 want the atomistic rigid bodies to adapt their orientation simultaneously. The fluctu-
1132 ations on the interactions due to the rotation of the molecule are expected to be
1133 much smaller than that due to the coarse-grained stretching modes. The proper
1134 back-coupling is added to the coarse-grained dynamics using

$$1135 \quad \quad \quad 1136 \quad \quad \quad 1137 \quad \quad \quad 1138 \quad \quad \quad f_1 \times r = T_1 = f_i \times u \quad (2-34)$$

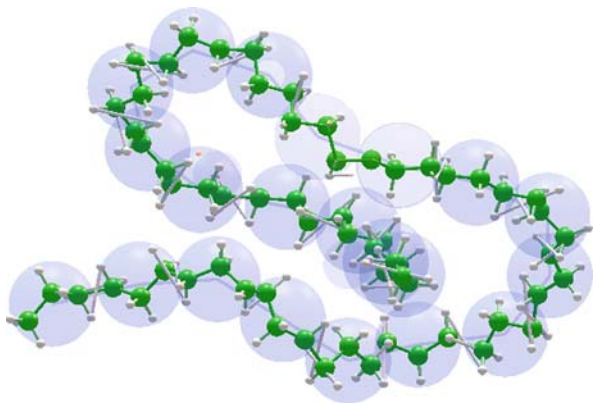
1139 That is, a back-coupling force f_1 on bead 1 is computed from the torque T_1 on the
1140 atomistic rigid body of this bead that is due to the atomistic interaction between one
1141 of its atoms i and another atom j belonging to neighboring bead 2 (see Figure 2-4) .
1142 In other words, the torque from the force f_i on atom i that drives the rotation results
1143 in addition to a force on the bead that is normal to the plane of the torque and the
1144 bond, r , between the beads.

1145 Having added the proper coupling between the rotational rigid atom dynamics
1146 and the rotational motions of the coarse-grained molecules (and removed the cou-
1147 pling with the translational bead motions) we obtain a stable energy conserving
1148 dynamics scheme which tends to thermal equilibrium between the two subsystems.
1149 In the following section, we will compare the two methods of on-the-fly reverse
1150 mapping through rigid rotation, and show that both the optimization scheme (or
1151 simulated annealing scheme) as well as the coupled dynamics scheme succeed
1152 in recovering good approximations of the atomistic structure in a coarse-grained
1153 simulation.

1154 2.3.4. Case Study 2: Polyethylene Chain

1157 In this case study, we compare the reverse mapping schemes, introduced in the pre-
1158 vious section, for coarse-grained simulations of a ($C_{75}H_{152}$) polyethylene chain in
1159 vacuum. The chain is represented by 25 coarse-grained beads, interacting through
1160 harmonic bond and bending potentials and Lennard–Jones type van der Waals inter-
1161 actions [53]. Each bead is mapped on 3 carbon atoms and its associated hydrogens
1162 as shown in Figure 2-5. We performed four NVT coarse-grained dynamics simu-
1163 lations of 500 ps in length at $T = 303$ K, each with a different variant of the reverse
1164 mapping scheme to recover atomistic configurations on the fly. The first 100 ps are
1165 regarded as equilibration; the remaining 400 ps of the trajectory was analyzed and
1166 compared with that of a fully atomistic simulation of the system. The four reverse
1167 mapping schemes are:

- 1169 1. At each timestep the relaxed atomistic C_3H_6 structure is inserted into each bead
1170 with a random orientation (i.e. random mapping).

*Multiscale Molecular Dynamics and the Reverse Mapping Problem*1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183

1184 *Figure 2-5.* Illustrative snapshot from a coarse-grained simulation of a polyethylene chain ($C_{75}H_{152}$),
1185 using the reverse mapping through rigid body rotation to recover atomistic configurations. Rotation
1186 dynamics of the atomistic C_3H_6 functions is governed by a simplified energy function containing only a
1187 harmonic C–C bond plus two repulsive H–H bonds (shown in *white*) between neighboring beads

1188
1189
1190
1191
1192
1193
1194
1195

2. The C_3H_6 structures are inserted at the beginning after which their orientations are optimized during the coarse-grained simulation through the damped $SO(3)$ rotational dynamics (i.e. damped rigid rotation).
3. Same as 2. but instead of annealing, the rotational dynamics is used, coupled to the coarse-grained dynamics (i.e. coupled rigid rotation).
4. atomistic flexible rotation

1196
1197
1198
1199

Here, the first scheme of random orientations is used for reference, the second and third schemes are the two rigid body rotation schemes introduced in the previous section, and the fourth scheme is another rotational scheme that acts on flexible atomistic structures, rather than rigid bodies.

1200
1201
1202
1203
1204
1205
1206
1207
1208

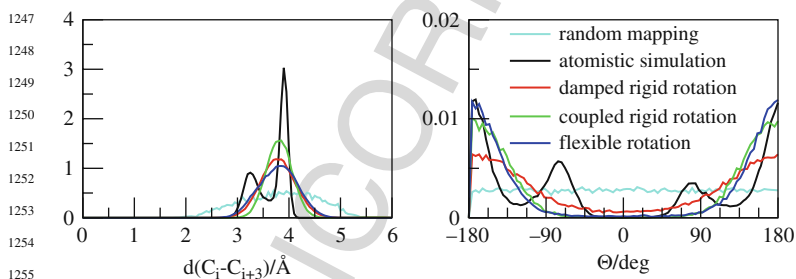
This last flexible scheme is added here because we will use it later in the hybrid multiscale method (see Section 2.1). That is, until now we have talked about reverse mapping to recover atomistic configurations from a coarse-grained trajectory, but hereafter, these reverse mapping schemes will be combined with the hybrid MD algorithm to pre-condition coarse-grained molecules before they enter the healing region. In order to gradually switch off this rotational dynamics while the atomistic interactions are being switched on, we require a variant of the rotational dynamics that will work on flexible, atomistically propagated, bodies as is the case in the healing region. This is the fourth scheme of which the details follow in Section 2.1.

1209
1210
1211
1212
1213
1214
1215

For the purpose of illustration, we used a very simple energy function in all of the three rotational dynamics schemes. That is, in principle we can use the full set of atomistic interactions of bonds, bends, torsions and even non-bonded potentials spanning neighboring beads, to govern the rotational dynamics, but it makes sense to choose an economic reduced subset instead. For our polyethylene chain, an intuitive good choice would include (at least) the carbon-carbon bond interaction spanning each pair of neighboring beads supplemented with the dihedral angles centered at

1216 this C–C bond. Here, in fact, we replace the dihedral angles with two repulsive
 1217 harmonic interactions between two hydrogens per pair of neighboring beads. These
 1218 repulsive “bonds” between hydrogens are shown in Figure 2-5 as white sticks.

1219 How do the reverse mapping schemes based on rotational dynamics of rigid or
 1220 flexible atomistic bodies perform with respect to the randomly oriented bodies and
 1221 with respect to a fully atomistic molecular dynamics simulation? To answer this
 1222 question, we compared atomistic distribution functions, two of which are shown
 1223 in Figure 2-6. On the left-hand side are the histograms of the pair-correlation of
 1224 each carbon in the chain with the carbon three positions away. This 1–4 correlation
 1225 is the shortest correlation that always spans two different beads. In the atomistic
 1226 simulation, the 1–4 correlation shows two peaks reflecting the staggered trans con-
 1227 figuration (larger peak) and the two, less favorable, staggered cis configurations
 1228 (smaller peak) of each quartet of carbons. In the coarse-grained simulations, the
 1229 repulsive hydrogen interactions (namely the white sticks in Figure 2-5) make the cis
 1230 configuration very improbable, so that the 1–4 correlations all show a single peak
 1231 positioned at the larger peak from the atomistic simulation. The broadening of the
 1232 peak from the reverse mapping schemes is due to the stretch vibration between the
 1233 beads, which is much softer than that between two carbons. The right-hand panel
 1234 in Figure 2-6 shows the histograms of dihedral angles between each sequence of
 1235 4 sequential carbon atoms centered at the bond between two beads (i.e. the first
 1236 two carbons belong to one bead and the second two carbons belong to the next
 1237 bead). Again, the repulsive terms in our simplified energy function used for the
 1238 rotational dynamics represses completely the secondary peaks seen in the histo-
 1239 grams from the atomistic simulation; however all three reverse mapping schemes
 1240 recover a good approximation to the average structure, which is the trans config-
 1241 uration. As a side remark, we note that the coupled dynamics scheme performs
 1242 slightly better, showing narrower distributions, than the damped dynamics scheme,
 1243 which is somewhat surprising as the damped dynamics would be expected to
 1244
 1245
 1246



1256 *Figure 2-6.* Comparison of the reverse mapping schemes to recover the atomistic structure in coarse-
 1257 grained simulation of a polyethylene chain. *Left panel:* histograms of the distance between each carbon at
 1258 position i with that at position $i + 3$ in the chain. *Right panel:* histograms of dihedral angles of sequential
 1259 carbon atoms of which the first two carbons and the second two carbons belong to different beads,
 1260 respectively

Multiscale Molecular Dynamics and the Reverse Mapping Problem

1261 tend closer to the lowest energy conformation. Further study to optimize the
1262 damping factor might improve this.

1263 All three reverse mapping schemes, based on rotational dynamics of rigid and
1264 flexible atomistic bodies, manage very well in maintaining good atomistic config-
1265 urations on the fly during a coarse-grained simulation. The overhead of the rigid
1266 body schemes is minimal due to the local nature of the energy function, relative
1267 to the demanding evaluation of the non-bonded interactions in the coarse-grained
1268 dynamics. The reverse mapping scheme was initially designed to allow for easy
1269 back-and-forth switching between atomistic and coarse-grained simulations of a
1270 system, reducing the requirement of re-equilibration in between. The reverse map-
1271 ping is therefore also expected to be helpful in replica exchange simulations, in
1272 which simultaneously several replicas of the system at different resolution are sim-
1273 ulated that can exchange based on a Metropolis criterion [18–21]. In the context of
1274 this chapter however, we will now return to the hybrid multiscale method and show
1275 how the rotational reverse mapping algorithm, applied to the coarse-grained region,
1276 is essential to simulate more complex structures.

1277

1278 **2.4. COMBINING ROTATIONAL REVERSE MAPPING** 1279 **WITH HYBRID MD**

1280

1281 In this last methodological section, we merge the reverse mapping algorithm with
1282 the adaptive multiscale method. The recent literature until now only shows applica-
1283 tions of adaptive hybrid atomistic/coarse-grained methods that consist of simple
1284 spherical beads in the low-resolution representation, for example, methane [7],
1285 water [54, 55], and even simpler toy models [8, 56]. The reason for this is that
1286 these spherical, or united, atoms require minor re-orientation when switching to the
1287 atomistic resolution compared to larger molecules spanning more than one coarse
1288 grained bead. In other words, all previous studies have explicitly avoided the reverse
1289 mapping problem. Indeed, it is clear from figure 4 that the dashed C–C bond imposes
1290 severe constraints on the relative orientation of the two beads representing a hexane
1291 molecule. If the atomic fragments in these two beads are randomly oriented with
1292 respect to one another when the hexane molecule enters the healing region, a very
1293 large healing region will be needed to bring the atomic representation down from an
1294 almost infinitely high potential energy value.

1295 Applying the reverse mapping algorithm in the coarse-grained region of the mul-
1296 tiscale setup allows, for the first time, hybrid MD simulations of realistic systems
1297 with multi-bead molecules in the low-resolution representation. The rigid body
1298 rotational dynamics (introduced in Section) maintains rigid atomistic structures
1299 superimposed on the coarse-grained molecular dynamics. These rigid structures are
1300 rotated about their centers of mass in accordance with an energy function which
1301 includes a subset of the local atomistic interactions, such as bond and bending
1302 potentials between atoms belong to sequential beads. This pre-conditioning of atom-
1303 istic configurations in the coarse-grained region therefore requires a much smaller
1304 healing region to equilibrate coarse-grained molecules that switch to the atomistic
1305 resolution.

1306 Note, however, that applying this reverse mapping scheme only to the coarse-
1307 grained region is not enough. Also at the coarse-grained end of the healing region,
1308 the orientation of atomistic structures within the beads becomes randomized when
1309 the molecules spend sufficient time at healing region positions of, say, around ninety
1310 percent coarse-grained character. Only a very wide healing region would then be
1311 able to restore the atomistic structure in a smooth manner, when such a molecule is
1312 pushed toward the atomistic region and the highly unfavorable atomistic interactions
1313 are switched on. This is the reason that a scheme in which atomistic structures are
1314 locally equilibrated (constrained to keep the centers of mass at the coarse-grained
1315 bead positions), for molecules that leave the coarse-grained region, does not suffice.
1316 Instead, we will also apply the rotational reverse mapping scheme in the heal-
1317 ing region, which is then gradually switched off, together with the coarse-grained
1318 interactions, as the atomistic forcefield takes over.

1319 In the healing region, in contrast to the coarse grained region, the atomistic bod-
1320 ies are not rigid. Flexible bodies break the $SO(3)$ rotational dynamics integrator
1321 because the moment of inertia tensor changes due to non-rotational forces. To never-
1322 theless allow for a rotational dynamics in conjunction with the atomistic dynamics,
1323 we employ a modified approach that adds the same reduced energy function that
1324 governs the rigid body rotational dynamics in the coarse-grained region. However,
1325 instead of computing a torque on the body from the atomistic interactions, we let
1326 the interaction act directly on the atoms with the additional constraint that the frag-
1327 ment should not feel an effective force. This constraint is satisfied if we apply a
1328 counter interaction on the fragment that cancels any translational acceleration and
1329 only keeps the rotational acceleration. In the example of Figure 2-4, this would
1330 mean that the atomistic bond interaction (dashed line) causes a force on the inter-
1331 acting carbon atoms and simultaneously a constraint force of the same amplitude but
1332 with opposite sign on the centers of mass. In the healing region, the coarse-grained
1333 forces acting on the centers of mass are distributed, mass-weighted, over the atoms.
1334 The coupling of this rotational dynamics with the normal atomistic dynamics in the
1335 healing region is implemented on the same footing as the coupling between those
1336 terms in the coarse-grained region (see Section 2.3.3). In Section , this atomistic
1337 flexible rotation variant of the reverse mapping technique was shown to behave well,
1338 compared to the rigid body rotational algorithms and to a fully atomistic molecular
1339 dynamics simulation.

1340

1341 **2.4.1. Case Study 3: Hybrid Simulation of a Polyethylene Chain**

1342

1343 In the third case study, we examine the application of the final combined adaptive
1344 multiscale molecular dynamics, i.e. including the reverse mapping algorithm, to the
1345 folding of a polyethylene chain in vacuum. The $C_{150}H_{302}$ chain is twice as long
1346 as in the previous case study and is represented by 50 beads at the coarse-grained
1347 level [53]. The initially stretched configuration is taken from an equilibration run,
1348 subject to an end-to-end distance constraint. As an illustration of the method, we
1349 show 2 short, 150 ps, hybrid MD simulations of the folding process, once using
1350 the reverse mapping algorithm and once without. Two atomistic regions are defined

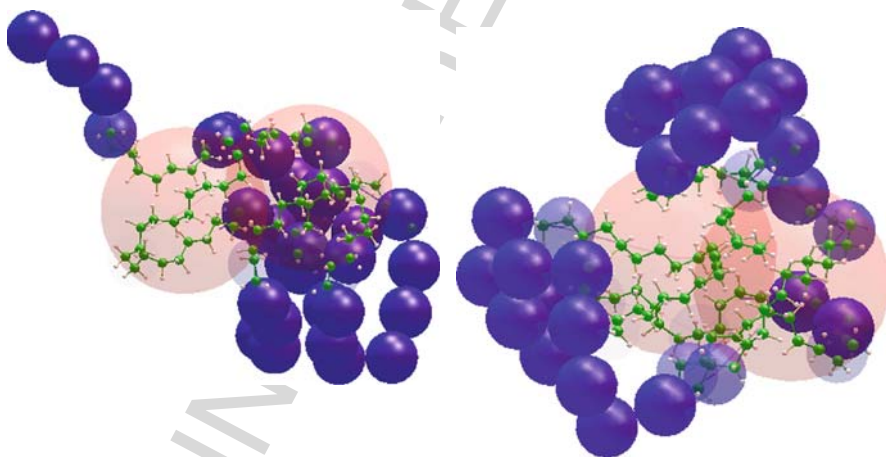
Multiscale Molecular Dynamics and the Reverse Mapping Problem

1351 with a radius of $R_A = 6 \text{ \AA}$ centered on beads 15 and 30 in the coarse-grained
1352 representation of the chains. The surrounding healing region skin is $R_{HR} = 5 \text{ \AA}$. Two
1353 representative snapshots of the hybrid MD simulation (using the reverse mapping
1354 algorithm) are shown in Figure 2-7.

1355 Comparing the total energies of the two simulations, with and without reverse
1356 mapping, displays a dramatic difference, as shown in Figure 2-8, top panel. Using
1357 the reverse mapping algorithm, which maintains a good pre-conditioned atomistic
1358 structure in the coarse-grained region, the simulation shows very good energy con-
1359 servation. Without the reverse mapping algorithm the total energy shows erratic
1360 behavior (note the difference in scales in these plots), indicating problems and poor
1361 accuracy in the simulation. Visual inspection of the trajectory shows that groups of
1362 atoms move suspiciously fast in the healing region and eventually, after about 95 ps
1363 of simulation, the system explodes. Because of the relatively large number of atoms
1364 in the healing region, the increasing temperature (despite the thermostat) is an indi-
1365 cation of problems (data not shown). Note that for example in a simulation of the
1366 chain in a solvent, such local temperature changes may not be apparent in the total
1367 system temperature.

1368 The bottom panel in Figure 2-8 shows the decreasing radius of gyration
1369 and the end-to-end distance of the polymer chain as it collapses from the initial
1370 stretched configuration to its random coil state. Note that, in this case, the
1371 unstable hybrid simulation displays reasonable behavior for these observables,
1372 not indicating any obvious problems, at least until the system explodes after
1373 95 ps.

1374 In conclusion, we have seen that incorporating the reverse mapping algorithm
1375 into our hybrid MD method to pre-condition the atomistic structure results in a
1376 robust adaptive multiscale molecular dynamics method. Close observation of the
1377



1393 *Figure 2-7.* Two snapshots from the adaptive multiscale simulation of a polyethylene chain ($C_{150}H_{302}$)
1394 in vacuum, using two spherical atomistic regions (*red spheres*). The atomistic regions have a diameter of
1395 6 \AA and are centered on beads 15 and 30 of the 50 bead coarse-grained representation (*blue spheres*)

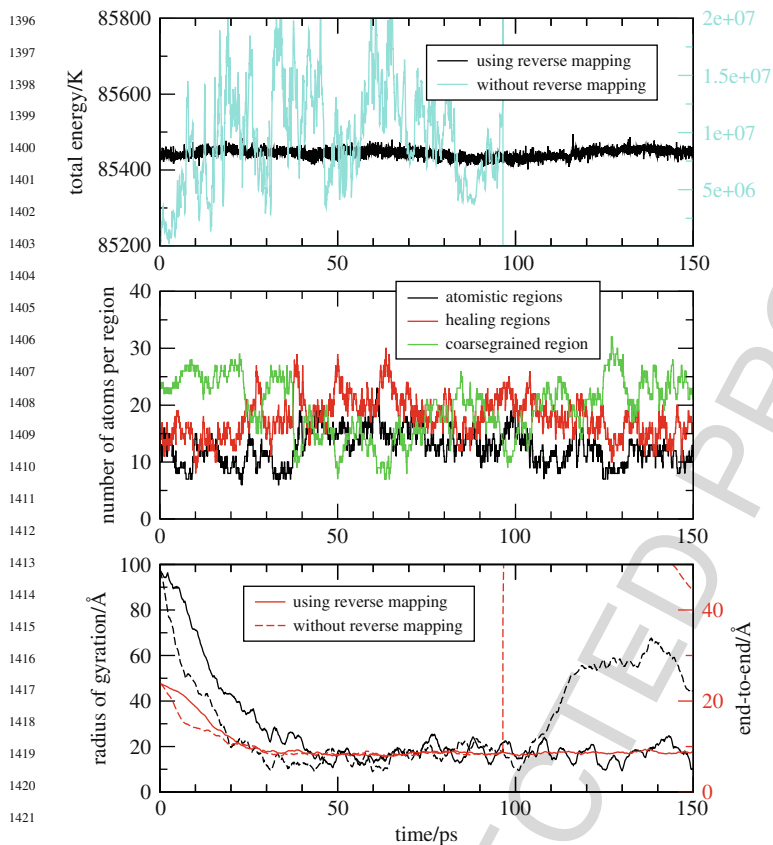


Figure 2-8. *Top panel:* Using the reverse mapping algorithm within the hybrid MD simulation of a polyethylene chain leads to good conservation of the total energy; instead without reverse mapping the hybrid method is unstable. *Middle panel:* the continuously fluctuating numbers of beads in the low, high, and hybrid resolution regions. *Bottom panel:* the folding of the chain shown by the radius of gyration (black lines; left axis) and the end-to-end distance (red lines; right axis). Despite the spurious dynamics of the simulation without reverse mapping, these observables show fortuitous similar physical behavior, that is, until the system explodes after about 95 ps

conservation of the total energy is required to assess the accuracy and physical relevance of the hybrid simulation. Without the reverse mapping algorithm, the hybrid method only works for small structureless molecules that can be represented by a single spherical bead in the coarse-grained region.

This case study of the collapse of a polymer chain illustrates the promising potential of the adaptive atomistic/coarse-grained method for the study of much more complicated and demanding phenomena of self-assembly, for example protein-protein interaction, multimeric protein assembly, and protein-DNA interaction. We foresee that such investigations, which now require either enormous computer

Multiscale Molecular Dynamics and the Reverse Mapping Problem

resources or rather simplified models, will take full advantage of the speedup from this multiscale method. Compared to a million atom sized MD simulation, such a speedup could easily reach two or three orders of magnitude when the demanding atomistic description can be limited to interacting regions of several hundreds to thousands of atoms.

1446

1447 **2.5. SUMMARY**

1448

Adaptive multiscale molecular dynamics is a promising new simulation technique aimed at bridging the gap between the large spatial and temporal scales exhibited by the phenomena that we wish to predict and the fast and short-ranged molecular fluctuations that limit current high-resolution methods. This technique allows one to focus the available computational resources on those special regions of the system where the key events are occurring by modeling the system in these regions at a higher, more accurate, resolution than the rest of the system. Here, we used an atomistic description in the high-resolution regions, and coarse-grained models, in which atoms are grouped into single interaction sites, to describe the other, low-resolution, regions.

The level of description of molecules that move between the high and low resolution regions adapts on the fly. This transformation from a coarse-grained description to an atomistic one, or vice versa, proceeds in a smooth manner through a thin intermediate *healing region* that bridges between the high and low resolution regions. In particular, the transformation from coarse-grained to atomistic, the so-called *reverse mapping*, is otherwise cumbersome as it requires the introduction of internal degrees of freedom that should be equilibrated together with their surrounding molecules. The continuous introduction (and removal) of degrees of freedom in the healing region is therefore a non-equilibrium process that produces heat, which can be removed with a thermostat. An important feature of the current multiscale algorithm is that it recovers, nevertheless, the total energy as a conserved quantity. *Energy conservation* provides a crucial handle to assess the accuracy of the integration of the equations of motion (i.e. the quality of the simulation) and is for example necessary to be able to choose an appropriate size for the healing region as well as for the time step.

We also discussed a special reverse mapping technique that allows one to obtain the atomistic conformations from a coarse-grained molecular dynamics simulation at low computational overhead. The algorithm consists of a dynamics on the Lie group $SO(3)$ of rotations for every coarse-grained site. Combining this technique with the adaptive multiscale approach to pre-condition molecules in the low-resolution region, before they enter the healing region, helps to limit the size of the healing region.

1481

1482 **ACKNOWLEDGMENTS**

1483

We are very grateful to Preston B. Moore, Peter G. Bolhuis, Michael L. Klein and Michele Parrinello for helpful discussions.

1485

AQ1

REFERENCES

- 1486 1. Wesolowski, T. A., Warshel, A., *J. Phys. Chem.* **1993**, *97*, 8050.
- 1487 2. Jacob, C., Neugebauer, J., Visscher, L., *J. Comput. Chem.* **2008**, *29*, 1011.
- 1488 3. Cramer, C. J., Truhlar, D. G., *Chem. Rev.* **1999**, *99*, 2161.
- 1489 4. Tomasi, J., Mennucci, B., Cammi, R., *Chem. Rev.* **2005**, *105*, 2999.
- 1490 5. Tomasi, J., Persico, M., *Chem. Rev.* **1994**, *94*, 2027.
- 1491 6. Gao, J., in *Reviews in Computational Chemistry*, edited by Lipkowitz, K., Boyd, D., Volume 7, pp.
1492 119–185, 119, VCH, New York, **1992**.
- 1493 7. Ensing, B., Nielsen, S. O., Moore, P. B., Klein, M. L., Parrinello, M., *J. Chem. Theory Comput.*
1494 **2007**, *3*, 1100.
- 1495 8. Praprotnik, M., Site, L. D., Kremer, K., *J. Chem. Phys.* **2005**, *123*, 224106.
- 1496 9. Heyden, A., Truhlar, D. G., *J. Chem. Theory Comput.* **2008**, *4*, 217.
- 1497 10. Shi, Q., Izvekov, S., Voth, G. A., *J. Phys. Chem. B* **2006**, *110*, 15045.
- 1498 11. Neri, M., Anselmi, C., Cascella, M., Maritan, A., Carloni, P., *Phys. Rev. Lett.* **2005**, *95*, 218102.
- 1499 12. Villa, E., Balaeff, A., Mahadevan, L., Schulten, K., *Multiscale Model. Simul.* **2004**, *2*, 527.
- 1500 13. Dupuy, L. M., Tadmor, E. B., Miller, R., Phillips, R., *Phys. Rev. Lett.* **2005**, *95*, 060202.
- 1501 14. Diestler, D. J., Zhou, H., Feng, R., Zeng, X. C., *J. Chem. Phys.* **2006**, *125*, 064705.
- 1502 15. Rudd, R. E., Broughton, J. Q., *Phys. Stat. Sol. B* **2000**, *217*, 251.
- 1503 16. Izvekov, S., Voth, G. A., *J. Phys. Chem. B* **2005**, *109*, 2469.
- 1504 17. Liu, P., Izvekov, S., Voth, G. A., *J. Phys. Chem. B* **2007**, *111*, 11566.
- 1505 18. Lyman, E., Ytreberg, F., Zuckerman, D., *Phys. Rev. Lett.* **2006**, *96*, 028105.
- 1506 19. Lyman, E., Zuckerman, D. M., *J. Chem. Theory Comput.* **2006**, *2*, 656.
- 1507 20. Christen, M., van Gunsteren, W., *J. Chem. Phys.* **2006**, *124*, 154106.
- 1508 21. Liu, P., Voth, G. A., *J. Chem. Phys.* **2007**, *126*, 045106.
- 1509 22. Csányi, G., Albaret, T., Payne, M. C., DeVita, A., *Phys. Rev. Lett.* **2004**, *93*, 175503.
- 1510 23. Hoogerbrugge, P. J., Koelman, J. M. V. A., *Europhys. Lett.* **1992**, *19*, 155.
- 1511 24. Smith, D. E., Harris, C. B., *J. Chem. Phys.* **1990**, *92*, 1304.
- 1512 25. Turq, P., Lantelme, F., Friedman, H. L., *J. Chem. Phys.* **1977**, *66*, 3039.
- 1513 26. Laio, A., Parrinello, M., *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562.
- 1514 27. Ensing, B., Vivo, M. D., Liu, Z. W., Moore, P. B., Klein, M. L., *Acc. Chem. Res.* **2006**, *39*, 73.
- 1515 28. Nielsen, S. O., Lopez, C. F., Srinivas, G., Klein, M. L., *J. Phys. Condens. Mater.* **2004**, *16*, R481.
- 1516 29. Henderson, R. L., *Phys. Lett. A* **1974**, *49A*, 197.
- 1517 30. Lyubartsev, A. P., Laaksonen, A., *Phys. Rev. E* **1995**, *52*, 3730.
- 1518 31. Soper, A. K., *Chem. Phys.* **1996**, *202*, 295.
- 1519 32. Akkermans, R., Briels, W., *J. Chem. Phys.* **2001**, *114*, 1020.
- 1520 33. Jain, S., Garde, S., Kumar, S. K., *Ind. Eng. Chem. Res.* **2006**, *45*, 5614.
- 1521 34. Njo, S. L., van Gunsteren, W. F., Mueller-Plathe, F., *J. Chem. Phys.* **1995**, *102*, 6199.
- 1522 35. Ercolessi, F., Adams, J. B., *Europhys. Lett.* **1994**, *26*, 583.
- 1523 36. Izvekov, S., Parrinello, M., Burnham, C. J., Voth, G. A., *J. Chem. Phys.* **2004**, *120*, 10896.
- 1524 37. Maurer, P., Laio, A., Hugosson, H. W., Colombo, M. C., Rothlisberger, U., *J. Chem. Theory Comput.*
1525 **2007**, *3*, 628.
- 1526 38. de Pablo, J. J., Curtin, W. A., *MRS Bull.* **2007**, *32*, 905.
- 1527 39. Nosé, S. J., *J. Chem. Phys.* **1984**, *81*, 511.
- 1528 40. Hoover, W. G., *Phys. Rev. A* **1985**, *31*, 1695.
- 1529 41. Parrinello, M., Rahman, A., *Phys. Rev. Lett.* **1980**, *45*, 1196.
- 1530 42. Bussi, G., Donadio, D., Parrinello, M., *J. Chem. Phys.* **2007**, *126*, 014101.

Multiscale Molecular Dynamics and the Reverse Mapping Problem

- 1531 43. MacKerel Jr., A., Brooks III, C., Nilsson, L., Roux, B., Won, Y., Karplus, M., CHARMM: The
1532 energy function and its parameterization with an overview of the program, in *The Encyclopedia of*
1533 *Computational Chemistry*, edited by v. R. Schleyer et al., P., Volume 1, pp. 271–277, John Wiley
1534 & Sons: Chichester, 1998.
- 1535 44. Jorgensen, W., Madura, J., Swenson, C., *J. Am. Chem. Soc.* **1984**, *106*, 6638.
- 1536 45. Abrams, C. F., *J. Chem. Phys.* **2005**, *123*, 234101.
- 1537 46. Praprotnik, M., Kremer, K., Site, L. D., *J. Phys. A Math. Theor.* **2007**, *40*, F281.
- 1538 47. Praprotnik, M., Kremer, K., Site, L. D., *Phys. Rev. E* **2007**, *75*, 017701.
- 1539 48. Heyden, A., Lin, H., Truhlar, D. G., *J. Phys. Chem. B* **2007**, *111*, 2231.
- 1540 49. Taylor, C. J., Kriegman, D. J., *IEEE Trans. Rob. Autom.* **1998**, *14*, 417.
- 1541 50. Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., *Numerical Recipes*, Cambridge
University Press: New York, **1992**.
- 1542 51. Nielsen, S. O., Ensing, B., Moore, P. B., Klein, M. L., Coarse grained to atomistic mapping
1543 algorithm: a tool for multiscale simulations, in *Multiscale simulation methods for nanomaterials*,
1544 edited by Ross, R., Mohanty, S., pp. 73–88, John Wiley and Sons, Inc.; Hoboken, NJ, **2008**.
- 1545 52. Krysl, P., Endres, L., *Int. J. Numer. Methods Eng.* **2005**, *62*, 2154.
- 1546 53. Nielsen, S. O., Lopez, C. F., Srinivas, G., Klein, M. L., *J. Chem. Phys.* **2003**, *119*, 7043.
- 1547 54. Praprotnik, M., Matysiak, S., Site, L. D., Kremer, K., Clementi, C., *J. Phys. Condens. Mater.* **2007**,
19, 292201.
- 1548 55. Matysiak, S., Clementi, C., Praprotnik, M., Kremer, K., Site, L. D., *J. Chem. Phys.* **2008**, *128*,
1549 024503.
- 1550 56. Praprotnik, M., Site, L. D., Kremer, K., *Phys. Rev. E* **2006**, *73*, 066701.
- 1551
- 1552
- 1553
- 1554
- 1555
- 1556
- 1557
- 1558
- 1559
- 1560
- 1561
- 1562
- 1563
- 1564
- 1565
- 1566
- 1567
- 1568
- 1569
- 1570
- 1571
- 1572
- 1573
- 1574
- 1575

1576 **CHAPTER 2**

1577

1578

1579

Q. No.**Query**

1580

1581

1582

1583

AQ1

We have changed the heading level from Bibliography to References in order to maintain book Consistency. Please check and confirm.

1584

1585

1586

1587

1588

1589

1590

1591

1592

1593

1594

1595

1596

1597

1598

1599

1600

1601

1602

1603

1604

1605

1606

1607

1608

1609

1610

1611

1612

1613

1614

1615

1616

1617

1618

1619

1620

UNCORRECTED PROOF