Benchmark of Schemes for Multiscale Molecular Dynamics Simulations

Goga, N.; Melo, M. N.; Rzepiela, A. J.; de Vries, Alex; Hadar, A.; Marrink, S. J.; Berendsen, Herman

Published in:
Journal of Chemical Theory and Computation

DOI:
10.1021/ct501102b

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Benchmark of Schemes for Multiscale Molecular Dynamics Simulations

N. Goga,[†,‡,§] M. N. Melo,[†,§] A. J. Rzepiela,[¶,†] A. H. de Vries,[†] A. Hadar,[‡] S. J. Marrink,*[,†] and H. J. C. Berendsen[†]

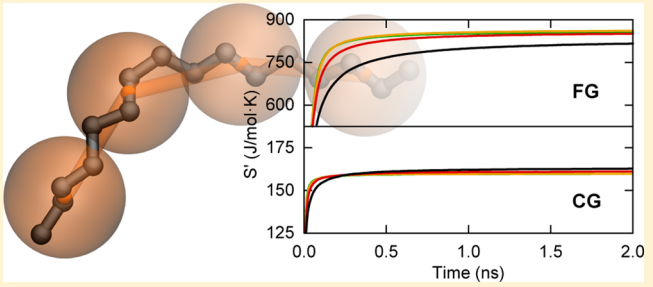
[†]Groningen Biomolecular Sciences and Biotechnology Institute, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, Groningen, The Netherlands

[‡]Politehnica University of Bucharest, Splaiul Independentei 313, 060042 Bucharest, Romania

[¶]Biozentrum, University of Basel, Klingelbergstrasse 50/70, CH-4056 Basel, Basel-Stadt, Switzerland

Supporting Information

ABSTRACT: In multiscale molecular dynamics simulations the accuracy of detailed models is combined with the efficiency of a reduced representation. For several applications — namely those of sampling enhancement — it is desirable to combine fine-grained (FG) and coarse-grained (CG) approaches into a single hybrid approach with an adjustable mixing parameter. We present a benchmark of three algorithms that use a mixing of the two representation layers using a Lagrangian formalism. The three algorithms use three different approaches for keeping the particles at the FG level of representation together: 1) addition of forces, 2) mass scaling, and 3) temperature scaling. The benchmark is applied to liquid hexadecane and includes an evaluation of the average configurational entropy of the FG and CG subsystems. The temperature-scaling scheme achieved a 3-fold sampling speedup with little deviation of FG properties. The addition-of-forces scheme kept FG properties the best but provided little sampling speedup. The mass-scaling scheme yielded a 5-fold speedup but deviated the most from FG properties.



1. INTRODUCTION

Molecular dynamics (MD) simulations give invaluable insights into the atomistic details and dynamics of molecular systems. Unfortunately the time scales and system sizes realistically within the reach of classical, all-atom MD simulations are rather limited. These limits can be eased at the expense of molecular detail by grouping a number of atoms together into coarse-grained particles.[1−3] Not surprisingly then, the main disadvantage of a coarse-grained (CG) model over a fine-grained (FG) model is exactly that the precise atomistic details are lost. In many applications it is important to preserve atomistic details for some region of special interest, while for the remainder of the system a CG approach suffices. In other applications it may be desirable to switch temporarily to a FG description when special events occur or when the validity of a CG treatment is to be assessed.

The question how to combine the strengths of both of these approaches in a so-called hybrid or multiscale MD simulation has received a great deal of attention[4−15] and is reviewed in refs 16−22. In the pioneering approaches of Christen and van Gunsteren[7] and Praprotnik et al.[8] all particles are represented both in fine-grained and in coarse-grained detail. These two representations are then coupled through a mixing parameter λ. The multiscale interaction forces are computed as a weighted λ sum of the interactions on FG and CG levels. Christen and van Gunsteren use a constant λ for the whole system, whereas Praprotnik et al. use a position-dependent λ.

If the λ-scaling is applied to FG bonded interactions then, in an almost fully CG system — i.e. $\lambda \approx 0$ — the bonded forces are not strong enough to keep bonded FG particles in their correct relative positions. This becomes problematic when switching to a stronger FG component (a higher λ) because the FG particles will have diffused away from each other, which then introduces artificially high forces in the system.

There are several possible representations of the mixed FG-CG system. We shall restrict ourselves in this article to the special case of constant (i.e., space- and time-independent) mixing parameter λ and to a Lagrangian description of the dynamics. The latter allows us to employ the rich machinery of statistical mechanics. In principle there are two choices to describe the CG particles in mixed systems: (i) CG particles are generated as *virtual* particles with neither mass nor kinetic energy; forces on CG particles are distributed over the FG particles that generate them. (ii) CG particles are dynamic particles with mass and kinetic energy; the relation between CG and FG positions are incorporated in the dynamics as a set of constraints. In the latter description the masses are scaled such

Received: December 5, 2014
Published: March 11, 2015

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

that the total mass remains independent of the mixing parameter. In these descriptions the use of *fractional degrees of freedom*[8] has been avoided. The former description is used in a modified form by Christen and van Gunsteren,[7] and the latter is used in the scaled Lagrangian method of Heyden and Truhlar.[11]

In this article we present a benchmarking of three multiscale algorithms — two adapted from the literature and one developed by us — corresponding to three possible approaches for stably simulating a multiscale system:

- *Addition of forces:* Christen and van Gunsteren[7] propose to scale all interactions at FG and CG levels, with the exception of the FG bonded interactions that are left entirely unscaled — a scheme termed 'multigraining'. The CG particles are virtual and massless. Since the FG bonded forces are not scaled, they will act on the molecules together with an increasing CG bonded contribution as λ approaches zero.
- *Temperature scaling:* We used a FG dynamic system with virtual CG particles, described by a system of generalized coordinates consisting of CG coordinates and *relative* FG coordinates, subject to constraints to satisfy the relations between CG and FG positions. Two subsystems (CG and FG) can be defined. In order for the FG subsystem to maintain a proper Boltzmann distribution, it must be kept at a λ-scaled temperature. As we will show, the Boltzmann distribution is preserved both at FG and CG levels. This is an algorithm developed by the authors of this paper.[20]
- *Mass scaling:* We used a mixed FG-CG system with scaled masses and constraints to satisfy the relations between CG and FG positions, valid for the special case of constant λ. This resembles (but is not equivalent to) the approach of Heyden and Truhlar[11] who consider a sum of weighted Lagrangians.

In all cases the potential is a linear combination of $(1-\lambda)$-scaled CG and λ-scaled FG components, with the exception that in the first case the bonded FG interactions are not scaled.

For the present article we restrict the treatment to systems with constant (i.e., space- and time-independent) λ. We focus on characterizing how the coupling between resolutions can enhance the sampling of the FG phase space. Systems at different λ values are tested, but sampling enhancement is assessed individually, without any exchange between systems. This is in contrast with proposed λ-based replica exchange methods.[6,7]

The benchmark is applied to two liquid hexadecane systems. A set of parameters developed to reproduce alkane surface tension[23] was used at the FG level for both systems. In one of the systems the CG level was simulated using the Martini force field[24] — a force field parametrized using empirical thermodynamic (partitioning) data. The CG force field for the second system was derived as a potential of mean force (PMF) from the underlying fine-grained model employing an iterative Boltzmann inversion (IBI) method.[25] Both CG models represent four consecutive methylene groups by one interaction site or bead.

2. MULTISCALING THEORY: TEMPERATURE SCALING

In the multiscaling theories that we are going to present we describe a system on two levels: FG particles (positions $\boldsymbol{r}$) with atomic detail and CG particles (positions $\boldsymbol{R}$) that represent a number of FG particles. Each CG coordinate vector $\boldsymbol{R}_i$ is determined by the center of mass (c.o.m.) of a number (this may differ per CG particle) of FG particles:

$$M_i\boldsymbol{R}_i = \sum_k m_{ik}\boldsymbol{r}_{ik} \tag{1}$$

$$M_i \stackrel{\text{def}}{=} \sum_k m_{ik} \tag{2}$$

The system moves under a linear combination of FG and CG potentials; the FG potential contributes a fraction λ, and the CG potential contributes a factor $1-\lambda$. For the time being the factor λ is a fixed constant.

2.1. Derivation of Forces. *2.1.1. Generalized Coordinates.* It is convenient to define generalized coordinates that include the CG coordinates $\boldsymbol{R}$. In this way there is only one system and the CG system is part of its description. As generalized coordinates we choose

a) the CG positions $\boldsymbol{R}_i$

b) the *relative* FG positions $\boldsymbol{s}_{ik} \stackrel{def}{=} \boldsymbol{r}_{ik}-\boldsymbol{R}_i$.

These generalized coordinates have $3N$ degrees of freedom (d.o.f.) *too many*, where N is the number of CG particles. However, these are compensated by the $3N$ *constraints* that follow from 1 and 2:

$$\boldsymbol{\sigma}_i(\boldsymbol{s}) \stackrel{\text{def}}{=} \sum_k m_{ik}\boldsymbol{s}_{ik} = \boldsymbol{0}, \quad i = 1, \ldots, N \tag{3}$$

The equations of motion must be solved, while these constraints remain satisfied. The constraints involve only the coordinates $\boldsymbol{s}$.

2.1.2. Lagrangian. In order to derive the equations of motion, we use the Lagrange formalism. The Lagrangian $\mathcal{L}$ is defined as a function of coordinates and their time derivatives and equals the kinetic energy (K) minus the potential energy (V):

$$\mathcal{L}(\boldsymbol{R}, \boldsymbol{s}, \dot{\boldsymbol{R}}, \dot{\boldsymbol{s}}) = K - V \tag{4}$$

Here, the kinetic energy is a function of velocities only

$$\begin{aligned} K(\dot{\boldsymbol{R}}, \dot{\boldsymbol{s}}) &= \sum_i \sum_k \frac{1}{2} m_{ik}\left(\dot{\boldsymbol{R}}_i + \dot{\boldsymbol{s}}_{ik}\right)^2 \\ &= \sum_i \frac{1}{2} M_i \dot{\boldsymbol{R}}_i^2 + \sum_i \sum_k \frac{1}{2} m_{ik}\dot{\boldsymbol{s}}_{ik}^2 \end{aligned} \tag{5}$$

(the cross products vanish as a result of 3). The kinetic energy is therefore the sum of the kinetic energy of the CG particles and the kinetic energy of the relative coordinates.

The potential energy is a function of coordinates only and is chosen as a mixture of CG and FG contributions:

$$V(\boldsymbol{R}, \boldsymbol{s}) = \lambda V^{\text{FG}}(\boldsymbol{R}, \boldsymbol{s}) + (1-\lambda)V^{\text{CG}}(\boldsymbol{R}) \tag{6}$$

This is the basic assumption. From here we wish to derive consistent equations of motion, assuring that the CG potential is thermodynamically consistent with the FG potential.

2.1.3. Generalized Momenta. In the Lagrangian formalism the generalized momentum, conjugate to a specific generalized coordinate, is the derivative of the Lagrangian with respect to the velocity of that coordinate. We denote the momentum conjugate to $\boldsymbol{R}_i$ with $\boldsymbol{P}_i$ and the momentum conjugate to $\boldsymbol{s}_{ik}$ with $\boldsymbol{p}_{ik}$.

$$\boldsymbol{P}_i = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{R}}_i} = M_i\dot{\boldsymbol{R}}_i \tag{7}$$

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

$$\boldsymbol{p}_{ik} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{s}}_{ik}} = m_{ik}\dot{\boldsymbol{s}}_{ik} \tag{8}$$

2.1.4. Equations of Motion. For a system with constraints, the Lagrangian equations of motion are (for any pair of conjugated coordinates and momenta q_i, p_i)

$$\dot{p}_i = \frac{\partial}{\partial q_i}[\mathcal{L}(q, \dot{q}) + \sum_n \boldsymbol{\mu}_n \cdot \boldsymbol{\sigma}_n(q)] \tag{9}$$

Here the index n numbers the constraints, and the coefficients $\boldsymbol{\mu}_n$ are Lagrange multipliers which follow from the constraint equations $\boldsymbol{\sigma}_n(q) = \mathbf{0}$. In our case the equations of motion become specifically

$$\dot{\boldsymbol{P}}_i = -\frac{\partial V(\boldsymbol{R}, \boldsymbol{s})}{\partial \boldsymbol{R}_i} \tag{10}$$

$$\dot{\boldsymbol{p}}_{ik} = -\frac{\partial V(\boldsymbol{R}, \boldsymbol{s})}{\partial \boldsymbol{s}_{ik}} + \boldsymbol{\mu}_i \frac{\partial}{\partial \boldsymbol{s}_{ik}} \sum_k m_{ik}\boldsymbol{s}_{ik} \tag{11}$$

Defining the CG and FG forces as follows

$$\boldsymbol{F}_i^{\mathrm{CG}} \overset{\mathrm{def}}{=} -\frac{\partial V^{\mathrm{CG}}(\boldsymbol{R})}{\partial \boldsymbol{R}_i} \tag{12}$$

$$\boldsymbol{F}_{ik}^{\mathrm{FG}} \overset{\mathrm{def}}{=} -\frac{\partial V^{\mathrm{FG}}(\boldsymbol{r})}{\partial \boldsymbol{s}_{ik}} \tag{13}$$

and using 7 and 8, 10 and 11 become

$$M_i\ddot{\boldsymbol{R}}_i = (1-\lambda)\boldsymbol{F}_i^{\mathrm{CG}} + \lambda \sum_k \boldsymbol{F}_{ik}^{\mathrm{FG}} \tag{14}$$

$$m_{ik}\ddot{\boldsymbol{s}}_{ik} = \lambda \boldsymbol{F}_{ik}^{\mathrm{FG}} + \boldsymbol{\mu}_i m_{ik} \tag{15}$$

These are the practical equations of motion. The Lagrange multipliers $\boldsymbol{\mu}_i$ (each of which are 3-D vectors) are to be determined such that the constraint eqs 3 remain satisfied. This means a simple *shift* of the individual accelerations $\ddot{\boldsymbol{s}}_{ik}$ by a constant vector $\boldsymbol{F}^{\mathrm{c}}{}_{ik}/m_{ik}$, thus compensating the displacement of the c.o.m. Here $\boldsymbol{F}^{\mathrm{c}}{}_{ik}$ is the *constraint force* acting on the (ik)th particle:

$$\boldsymbol{F}_{ik}^{\mathrm{c}} = -\lambda \frac{m_{ik}}{M_i} \sum_k \boldsymbol{F}_{ik}^{\mathrm{FG}} \tag{16}$$

Alternatively, in the spirit of SHAKE,[26] one may in the Verlet algorithm first compute the displacements without taking the constraints into account, yielding coordinates $\boldsymbol{s}'_{ik}$, and then displace each of the coordinates with a vector such that 3 is satisfied:

$$\boldsymbol{s}_{ik} = \boldsymbol{s}'_{ik} + \Delta \boldsymbol{s}_i \tag{17}$$

with

$$\Delta \boldsymbol{s}_i = -\frac{\sum_k m_{ik}\boldsymbol{s}'_{ik}}{M_i} \tag{18}$$

The latter method is more robust, as the constraints remain exactly satisfied. Applying constraint forces as calculated from 16, the c.o.m. of the FG particles may slowly drift away from the CG position.

2.2. Temperature. *2.2.1. FG and CG Temperatures.* The system contains two subsystems: the CG positions $\boldsymbol{R}$ with velocities $\dot{\boldsymbol{R}}$ and the relative FG positions $\boldsymbol{s}$ with velocities $\dot{\boldsymbol{s}}$. If the coupling between the two systems is not very strong, the thermal equilibration within each subsystem may well be faster than the thermal exchange between the subsystems. In that case it makes sense to speak of two separate *temperatures*: T^{FG} for the FG subsystem and T^{CG} for the CG subsystem. One may artificially maintain two different temperatures by coupling the subsystems to two different temperature baths. The situation is somewhat similar to the ab initio MD method of Car and Parrinello,[27] where the nuclear motions have a temperature T, while the subsystem of electronic wave functions is simultaneously maintained at a very low temperature.

The two temperatures are defined by the kinetic energy and the number of degrees of freedom of each subsystem. For the CG subsystem with N particles there are $3N$ d.o.f., and the temperature is given by

$$T^{\mathrm{CG}} = \frac{\sum_i M_i(\dot{\boldsymbol{R}}_i)^2}{3Nk_{\mathrm{B}}} \tag{19}$$

where k_B is the Boltzmann constant. The FG subsystem has $3(n-N)$ d.o.f. when there are n FG particles. Hence the temperature is given by

$$T^{\mathrm{FG}} = \frac{\sum_i \sum_k m_{ik}(\dot{\boldsymbol{s}}_{ik})^2}{3(n-N)k_{\mathrm{B}}} \tag{20}$$

2.2.2. Statistical-Mechanical Considerations. Ideally, a CG description of a system preserves the thermodynamic quantities of the FG description. In particular, for a well-designed CG potential, the distribution of equilibrium configurations $w(\boldsymbol{R})$ of CG-positions should be the same for the full FG system ($\lambda = 1$) and the full CG system ($\lambda = 0$):

$$w^{\mathrm{CG}}(\boldsymbol{R}) \propto \exp[-\beta V^{\mathrm{CG}}(\boldsymbol{R})] \tag{21}$$

$$w^{\mathrm{FG}}(\boldsymbol{R}) \propto \int \exp[-\beta V^{\mathrm{FG}}(\boldsymbol{R}, \boldsymbol{s})]\, \delta(\sigma)\, \mathrm{d}\boldsymbol{s} \tag{22}$$

$$w^{\mathrm{CG}}(\boldsymbol{R}) \approx w^{\mathrm{FG}}(\boldsymbol{R}) \tag{23}$$

In 22, $\delta(\sigma)$ is a shorthand notation for a product of delta functions

$$\delta(\sigma) = \Pi_i \delta(\boldsymbol{\sigma}_i(\boldsymbol{s})) \tag{24}$$

that takes care of the constraints in the integration over all allowed $\boldsymbol{s}$ coordinates.

Note that 21−23 imply that

$$V^{\mathrm{CG}}(\boldsymbol{R}) = -k_{\mathrm{B}}T \ln \int \exp[-\beta V^{\mathrm{FG}}(\boldsymbol{R}, \boldsymbol{s})]\, \delta(\sigma)\, \mathrm{d}\boldsymbol{s} + const. \tag{25}$$

that is, the CG potential is a *potential of mean force* with respect to the $\boldsymbol{s}$ coordinates. In practice, since the CG potential is not usually derived from FG simulations and a simplified description is used, the actual CG potential may not exactly fulfill 25. In that case also 23 is not exactly fulfilled.

Now, what will be the distribution function of $\boldsymbol{R}$, generated by a mixed system with parameter λ? Assuming that the FG subsystem is equilibrated at a temperature $T^{\mathrm{FG}} = (k_{\mathrm{B}}\beta^{\mathrm{FG}})^{-1}$, this distribution will be

$$w^{\lambda}(\boldsymbol{R}) \propto \int \exp[-\beta^{\mathrm{FG}}\lambda V^{\mathrm{FG}}(\boldsymbol{R}, \boldsymbol{s})]\, \delta(\sigma)\, \mathrm{d}\boldsymbol{s} \tag{26}$$

It is clear that this distribution is compatible with 22 only if $\beta^{\mathrm{FG}}\lambda = \beta$ meaning that the FG temperature must be

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

$$T^{\mathrm{FG}} = \lambda T \tag{27}$$

This is not surprising, since we wish the distribution of FG particles to be given by the FG potential, irrespective of λ. Thus, if we scale V^{FG}, we should also scale T^{FG} to keep the Boltzmann factor the same. Conversely, if we would scale the FG potential without scaling the temperature, the FG subsystem would tend to an ideal gas at low values of λ, with a behavior that would deviate radically from the required distribution and with the FG particles flying wildly throughout the system.

3. MULTISCALING THEORY: MASS SCALING

One can choose to scale the masses in addition to scaling the energy function. In the theory presented in the previous section masses are not scaled; this leads to a FG subsystem that must be maintained at a scaled temperature in order to preserve the Boltzmann distribution. In this version we choose to scale the masses instead. Now there is only one temperature for the system as a whole. The total energy is conserved.

3.1. Lagrangian and Equations of Motion. The potential energy is defined in 6 in the same λ-dependent way as for temperature scaling. Likewise, we define generalized coordinates and derive the equations of motion from a Lagrangian extended with constraint conditions. There is no need to define relative coordinates, and the generalized coordinates are the CG coordinates $\boldsymbol{R}$ and the FG coordinates $\boldsymbol{r}$. The masses of CG and FG particles are scaled to $(1-\lambda)M_i$ and λm_{ik}, respectively. The constraints are given by 1, which can be written in the form

$$\boldsymbol{\sigma}_i = M_i\boldsymbol{R}_i - \sum_k m_{ik}\boldsymbol{r}_{ik} = \boldsymbol{0} \tag{28}$$

Note that in the constraint definition the masses are not scaled: for all λ the CG positions remain at the c.o.m. position of the corresponding FG particles. The Lagrangian, augmented with Lagrange multipliers $\boldsymbol{\mu}$ to satisfy the constraints, now reads

$$\mathcal{L}(\boldsymbol{R}, \boldsymbol{r}, \dot{\boldsymbol{R}}, \dot{\boldsymbol{r}}) = K - V + \sum_i \boldsymbol{\mu}_i\cdot\boldsymbol{\sigma}_i \tag{29}$$

We first obtain the generalized momenta $\boldsymbol{P}_i = (1-\lambda)M_i\dot{\boldsymbol{R}}_i$ and $\boldsymbol{p}_{ik} = \lambda m_{ik}\dot{\boldsymbol{r}}_{ik}$ as derivatives of $\mathcal{L}$ to the velocities. Next we find the equations of motion from the time derivatives of the generalized momenta as derivatives of $\mathcal{L}$ to the coordinates. These equations of motion contain derivatives of the constraint terms in 29. Solving for the constraints in 28 we obtain

$$\ddot{\boldsymbol{R}}_i = \frac{1}{M_i}\boldsymbol{F}_i^{\mathrm{CG}} - \lambda\boldsymbol{a}_i^{\mathrm{c}} \tag{30}$$

$$\ddot{\boldsymbol{r}}_{ik} = \frac{1}{m_{ik}}\boldsymbol{F}_{ik}^{\mathrm{FG}} + (1-\lambda)\boldsymbol{a}_i^{\mathrm{c}} \tag{31}$$

where

$$\boldsymbol{a}_i^{\mathrm{c}} = \frac{1}{M_i}\left(\boldsymbol{F}_i^{\mathrm{CG}} - \sum_k \boldsymbol{F}_{ik}^{\mathrm{FG}}\right) \tag{32}$$

Note that the constraint shift is scaled. When $\lambda = 1$ (complete FG), the CG position is drawn to the FG c.o.m.; when $\lambda = 0$ (complete CG), the FG c.o.m. is drawn to the CG position.

Both the FG and CG coordinates evolve in time as they would in a full FG or CG simulation; the only extra accelerations, $\boldsymbol{a}_i^{\mathrm{c}}$, come from the constraints. This means a simple *shift* of the individual accelerations $\dot{\boldsymbol{R}}_i$ and $\dot{\boldsymbol{r}}_{ik}$ by a constant vector proportional to a *constraint acceleration*, thus compensating the displacement of the c.o.m. It can also be seen that the *average* constraint acceleration is zero when the CG potential is the exact potential of mean force. If the constraint acceleration is monitored during a mixed simulation, the average can be used to adjust parameters of the CG potential.

3.2. Generated Ensemble, Energy, and Temperature. As the system is Hamiltonian, it will follow the canonical distribution. However, one has to be careful when constraints are involved, because the average obtained in constrained simulations needs a so-called *metric tensor correction*. This works out as an extra weight factor $|Z|^{-1/2}$ in configuration space, where Z is the following matrix:[28]

$$Z_{\alpha\beta} = \sum_j \frac{1}{\lambda m_j}\frac{\partial\sigma_\alpha}{\partial\boldsymbol{r}_j}\cdot\frac{\partial\sigma_\beta}{\partial\boldsymbol{r}_j} \tag{33}$$

Here j runs over all particles of the system, and σ_α is the αth constraint equation. When we evaluate Z, it turns out to be a diagonal matrix with constant values ($Z_{ii} = \lambda M_i$). Hence its determinant is a constant, and we can conclude that there is no configuration-dependent weight factor involved. All computed averages are correct canonical averages.

So, the probability of a configuration generated by constrained MD will be given by

$$\begin{aligned} w(\boldsymbol{R}, \boldsymbol{r}) &\propto \exp[-\beta\{(1-\lambda)V^{\mathrm{CG}}(\boldsymbol{R}) + \lambda V^{\mathrm{FG}}(\boldsymbol{r})\}] \\ &= \exp[-\beta(1-\lambda)V^{\mathrm{CG}}(\boldsymbol{R})]\exp[-\beta\lambda V^{\mathrm{FG}}(\boldsymbol{r})] \end{aligned} \tag{34}$$

subject to the constraint conditions 28. This equation suggests that the distribution can be factorized into two independent distributions but that is not true because the $\boldsymbol{R}$- and $\boldsymbol{r}$-spaces are not independently sampled. Dynamics samples only the hyperspace defined by the constraints.

The *temperature* is given by equipartition over the $3n$ d.o.f. of the system (n = number of FG particles)

$$K = \frac{3}{2}nk_{\mathrm{B}}T \tag{35}$$

where K is the total kinetic energy with scaled masses. In practical simulations the overall c.o.m. motion, which represents three uncoupled d.o.f., is kept to zero, and n should be replaced by $n-1$. The *total energy* $K+V$ should be conserved in the simulation.

3.3. Statistical-Mechanical Considerations. We now consider the equilibrium distribution functions of the CG coordinates $\boldsymbol{R}$ in the mixed case and compare these with the full CG system. The distribution function of $\boldsymbol{R}$ generated by a mixed system with parameter λ can be found by integrating 34 over $\boldsymbol{r}$:

$$w^{\lambda}(\boldsymbol{R}) \propto \int \exp[-\beta V(\boldsymbol{R}, \boldsymbol{r})]\delta(\sigma)\mathrm{d}\boldsymbol{r} \tag{36}$$

$$= \exp[-\beta(1-\lambda)V^{\mathrm{CG}}(\boldsymbol{R})]\int \exp[-\beta\lambda V^{\mathrm{FG}}(\boldsymbol{r})]\delta(\sigma)d\boldsymbol{r} \tag{37}$$

Here

$$\delta(\sigma) = \Pi_i\delta(\boldsymbol{\sigma}_i(\boldsymbol{s})) \tag{38}$$

restricting the integrand to the constraint hypersurface.

Now, if

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

$$\int \exp[-\beta\lambda V^{\mathrm{FG}}(\boldsymbol{r})]\,\delta(\sigma)\,d\boldsymbol{r} \propto \exp[-\beta\lambda V^{\mathrm{CG}}(\boldsymbol{R})] \tag{39}$$

then

$$w^{\lambda}(\boldsymbol{R}) \propto \exp[-\beta V^{\mathrm{CG}}(\boldsymbol{R})] \tag{40}$$

which is the expected behavior for a perfect potential of mean force at the temperature of the simulation. Note, however, that 39 represents the potential of mean force PMF for a given λ, which is not the same as the PMF for the full CG potential. Therefore, 40 is not equivalent to 37; equality would imply that the PMF is independent of temperature.

4. MULTISCALING THEORY: FORCE ADDITION

The force addition, or multigraining, scheme was proposed by Christen and van Gunsteren.[7] In that work it was coupled to a replica exchange process as had been described earlier.[6] Here we disregard the replica exchange application and focus on the force addition CG/FG multiscale coupling. In order to keep the FG particles together and preserve the temperature, bonded forces are kept unscaled with λ. We will rewrite the formalism to follow a pattern similar to the ones of the previous sections. Note that our λ (measuring the FG contribution) is equal to $(1-\lambda)$ used by Christen and van Gunsteren;[7] their λ measures the CG contribution.

4.1. Equations of Motion. The system consists of FG particles, numbered ik with masses m_{ik} and positions $\boldsymbol{r}_{ik}$ with atomic detail. CG positions are reconstructed from the c.o.m. of the respective FG atoms, and CG velocities are derived from FG ones.

The equations of motion concern only the FG particles and follow Hamiltonian mechanics with a contribution to the potential arising from the CG potential. Thus,

$$V(\boldsymbol{r}) \overset{\mathrm{def}}{=} (1-\lambda)V^{\mathrm{CG}}(\boldsymbol{R}(\boldsymbol{r})) + \lambda V^{\mathrm{nonb,FG}}(\boldsymbol{r}) + V^{\mathrm{bonded,FG}}(\boldsymbol{r}) \tag{41}$$

and the equations of motion become specifically

$$m_{ik}\ddot{\boldsymbol{r}}_{ik} = \boldsymbol{F}_{ik}^{\mathrm{bonded,FG}} + \lambda\boldsymbol{F}_{ik}^{\mathrm{nonb,FG}} + \frac{\boldsymbol{m}_{ik}}{\boldsymbol{M}_i}(1-\lambda)\boldsymbol{F}_i^{\mathrm{CG}} \tag{42}$$

where $\boldsymbol{M}_i$ and $\boldsymbol{F}^{\mathrm{CG}}{}_i$ are the mass and the force on the corresponding CG virtual particle.

It should be noted from 42 that bonded FG forces are present even if the multiscaling simulation is fully CG, at $\lambda = 0$. The presence of these forces ensures that the FG particles are kept together even at low values of λ, and there is no need for scaling the masses or the temperature as presented in the previous sections.

5. STOCHASTIC THERMOSTATS

In our benchmarks using IBI potentials, when the multiscaling system is close to the CG representation (λ values close to 0), a phenomenon similar to a *flying ice cube* effect[29] occurs for the temperature-scaling and force addition algorithms due to heat flow between CG and FG representations. In the flying ice cube effect the thermostat produces a noise that introduces a small heat production in the system, which accumulates especially in those d.o.f. that are uncoupled, or very weakly coupled, to the rest of the system — i.e., the overall translation and rotation of (part of) the system. A global thermostat cools the whole system down, distributing its heat flow over all d.o.f., the result being a cold system with too much energy in the overall translation and rotation. For counteracting such effects we extended the Berendsen thermostat with a weak stochastic component in two different ways, both described in detail in ref 30: the first thermostatization approach was the use of a Stochastic Dynamics (SD) extension to the Berendsen thermostat, implemented as an impulsive Langevin leapfrog algorithm. The second extension to the Berendsen thermostat involves the use of isotropic Dissipative Particle Dynamics (DPD-ISO), where the friction and noise are applied between pairs of particles. The DPD-ISO approach conserves the total momentum of the particles, in contrast with the Berendsen-SD thermostat. It is also implemented so that the stochastic heat exchange is local to each particle, which might further mitigate the formation of flying ice cubes. It should be noted that in its current implementation the DPD-ISO thermostat cannot be used with systems where bond constraints are used. However, a rather simple but exact modification is now available[31] that allows DPD-like thermostatting for systems with constraints.

6. ITERATIVE BOLTZMANN INVERSION

In structure-based coarse graining, a CG potential is determined in such a way that predefined target functions, which structurally characterize the system, are reproduced in the CG simulation. In the iterative Boltzmann inversion method[25] radial distribution functions (RDFs), $g_{ref}(r)$, are the target functions. The $g_{ref}(r)$ can be obtained either from experiment or from atomistic simulations. Through the simple Boltzmann inversion

$$V_{\mathrm{PMF}}(r) = -k_{\mathrm{B}}T\,\ln g_{ref}(r) \tag{43}$$

where k_{B} denotes the Boltzmann constant and T the temperature, the potential of mean force V_{PMF} between pairs of CG particles can be obtained. Unfortunately, this PMF cannot be directly used in a CG model because part of it represents multibody contributions from all the particles in the system. An iterative procedure should therefore be used:

$$V_{i+1}^{\mathrm{CG}}(r) = V_i^{\mathrm{CG}} + k_{\mathrm{B}}T\,\ln\left[\frac{g_i(r)}{g_{ref}(r)}\right] \tag{44}$$

The procedure is initiated with the V_{PMF} potential from the simple Boltzmann inversion. The subscript i denotes the iteration number. According to the Henderson[32] theorem the IBI method should give a unique two-body CG interaction for the given $g_{ref}(r)$. However, in practice convergence is difficult to achieve and varies for different parts of the potential.[33,34] Especially attractive parts require an extensive number of iterations, whereas repulsive parts converge relatively fast. Nevertheless, $g_{ref}(r)$ is reproduced with a good accuracy already after a small number of iterations. To enhance the convergence during the iteration procedure the CG potential can be modified to reproduce the target pressure. For example, a correction to the potential scaling linearly with the pressure difference can be applied[35,36]

$$\Delta V(r) = A\left(1 - \frac{r}{R_c}\right) \tag{45}$$

where R_c is a user defined cut off for the pair interaction potential, and A can be defined as

$$A = -|(\Delta P)|k_{\mathrm{B}}Tf \tag{46}$$

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

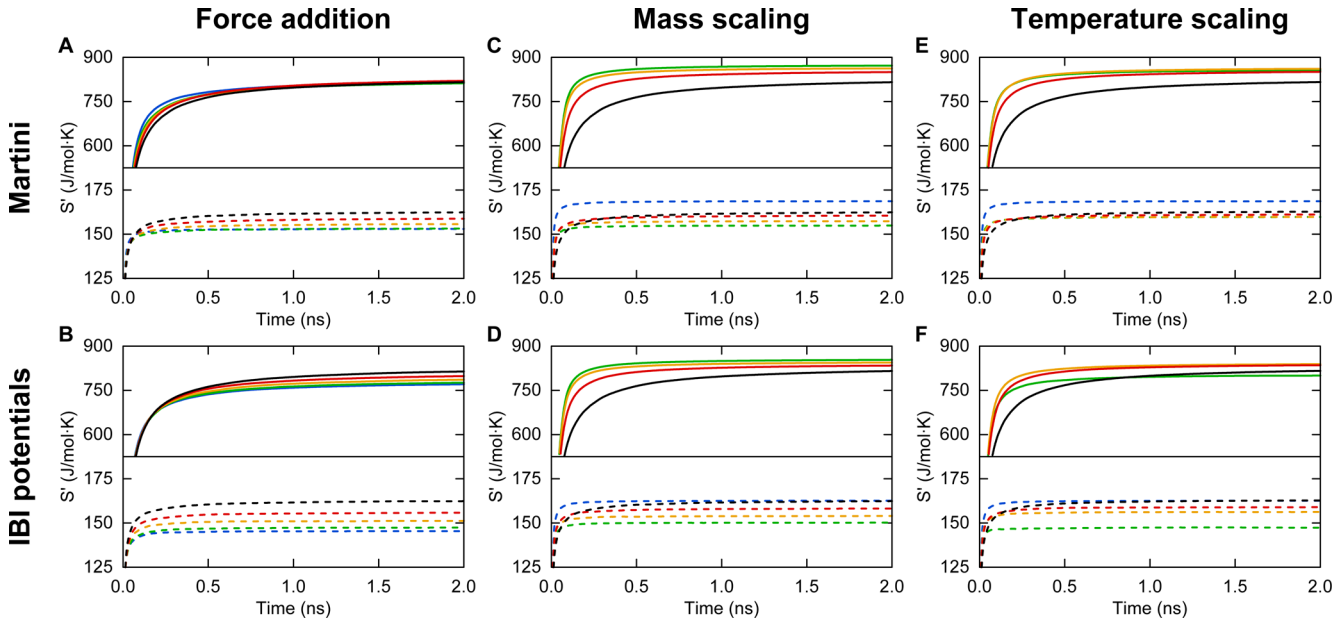


Figure 1. Comparison of FG and CG entropy buildup. Different schemes are compared: **A** and **B** — Force addition using the Martini (**A**) or IBI potentials (**B**); **C** and **D** — Mass scaling using the Martini (**C**) or IBI potentials (**D**); **E** and **F** — Temperature scaling using the Martini (**E**) or IBI potentials (**F**). Solid and dashed lines represent FG and CG buildup, respectively. Lines are colored according to the λ value: $\lambda = 1$, black; $\lambda = 0.75$, red; $\lambda = 0.5$, yellow; $\lambda = 0.25$, green; and $\lambda = 0$, blue. The curves at $\lambda = 1$ for the mass scaling scheme were not determined and the ones from panel **A** were used instead in panels **C** and **D**. The curves at $\lambda = 0$ were also not obtained for the temperature and mass scaling schemes; CG entropy buildup curves from pure Martini or IBI potential simulations were used in panels **C**, **D**, **E**, and **F**.

with ΔP being the pressure mismatch and f being a scaling factor.

7. SIMULATION AND ANALYSIS DETAILS

7.1. System Setup and Simulation. We investigated the three multiscaling algorithms on two hexadecane systems, which differ at the CG level. Both systems consisted of 320 hexadecane molecules. For both CG subsystems a 4-to-1 mapping was used, where the center of mass of four consecutive methylene groups is represented by a single particle or bead. For one of the systems the Martini CG force field[24] was used to describe the bonded and nonbonded interactions between CG particles. For the other system we derived the CG force field as a potential of mean force from the underlying fine grained model through the IBI method described earlier.[25]

Five λ values were studied (0, 0.25, 0.5, 0.75, and 1). The temperature scaling scheme becomes undefined at $\lambda = 0$, and the mass scaling one at $\lambda = 0$ and $\lambda = 1$. For these cases reference values from pure CG ($\lambda = 0$) and FG ($\lambda = 1$) simulations are presented for comparison, where appropriate. The simulations were performed using our modified version 4.0 of the GROMACS software package or the 4.5.5 version (unmodified) for the force addition scheme.[37] A united-atom representation was used for the FG hexadecane subsystem, with interaction parameters taken from ref 23. Newton's equations of motion were integrated using the leapfrog algorithm[38] with a 2 fs time step at both levels. All nonbonded interactions within a 1.4 nm short-range cutoff were evaluated every time step, based on a pair list recalculated every 5 steps. Note that in either subsystem the electrostatic interactions are zero. At the CG level, for the Martini system, a Lennard-Jones potential shifted between 0.9 and 1.2 nm was used to calculate nonbonded interactions (intramolecularly, beyond nearest bonded neighbors). Constraints were not used at either level of description. The forces and positions of the CG beads and of the FG atoms were calculated for each algorithm as described in the previous sections. The simulations were carried out in a cubic box of dimension 5.27 nm with an isothermal (NVT) ensemble, using a Berendsen thermostat[39] — or the Berendsen-SD and Berendsen-DPD extensions to it — with a coupling constant $\tau_T = 0.01$ ps. At the CG level, the temperature was always coupled to a bath at a reference temperature of 300 K, except in the case of the force addition scheme, where no CG thermostatization was applied. For analysis, the atomic coordinates were saved every 2 ps.

See the Supporting Information for the used inverted potentials and the structural and entropy-convergence data for all the scheme/potential combinations.

7.2. Data Analysis. The MD data were analyzed using standard and adapted GROMACS tools. Structural data (angle and dihedral distributions and c.o.m.-to-c.o.m. intermolecular radial distribution functions) is presented below for selected scheme/CG-potential combinations, but the results for all the systems can be found in the Supporting Information.

In order to investigate whether the mixed systems sample a similar configuration space as the pure CG or FG systems, the intramolecular configurational entropy was measured. This entropy can be determined approximately from the covariance matrix of the distribution of internal coordinates,[40] either for the CG coordinates (for all $\lambda \in [0,1]$) or for the FG coordinates (for $\lambda > 0$). When the sampled space for two simulations is comparable, at least their configurational entropies should be similar. A large discrepancy indicates a large difference in probability distribution and therefore unreliable thermodynamic properties.

The configurational entropy of the chains was calculated at both the FG and the CG level from the mass-weighted covariance matrix $\mathbb{D}$ in the quasi-harmonic approximation due to Schlitter:[40]

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

Table 1. Speedup in the Time Required for Reaching 98% of the Converged FG Entropy Buildup, Relative to the Full FG Case ($\lambda = 1$)

	force addition		mass scaling		temperature scaling	
	Martini	IBI potentials	Martini[a]	IBI potentials[a]	Martini	IBI potentials
$\lambda = 0$	1.40	1.34				
$\lambda = 0.25$	1.23	1.41	4.91	5.10	3.59	3.64
$\lambda = 0.5$	1.17	1.30	4.38	4.56	3.22	3.73
$\lambda = 0.75$	1.22	1.15	2.60	2.75	2.17	2.57

[a]Speedup factors relative to the $\lambda = 1$ curve of the force addition scheme.

$$S_{\text{true}} \leq S' = \frac{k_B}{2} \ln \det\left(\mathbb{I} + \frac{k_B T e^2}{\hbar^2}\mathbb{D}\right) \quad (47)$$

Here, S' is the calculated upper bound to the true configurational entropy S_{true}, $\mathbb{I}$ is the unit matrix, e is the base to the natural logarithm, and $\hbar$ is Planck's constant divided by 2π. Before calculation of S', overall translation and rotation of the molecules during the simulation is removed by shifting the center of mass of each configuration to the origin of the simulation volume and performing a best rotational fit of the atoms to the average structure. This procedure and application to FG and CG hexadecane has been described previously.[41] For all schemes, including mass and temperature scaling, the unscaled masses and temperatures were used to evaluate 47.

8. RESULTS

8.1. Entropy Buildup. Figure 1 represents the buildup of the average configurational entropy of the FG and CG subsystems, as a function of time, for the different schemes and potentials. In agreement with previous results, the configurational entropies are nearly converged after 10 ns of simulation. The buildup curves are similar to the ones reported previously at full FG and CG descriptions,[41] the final values agreeing to within a few percent. Note that the absolute values of the configurational entropies of FG and CG descriptions are different because the number of atoms (16 per molecule at the FG level) and beads (4 per molecule at the CG level) from which the number is calculated differs.

Table 1 lists, for each scheme and λ value, the speedup relative to a pure FG simulation in reaching a configurational entropy 98% of the converged value. Introducing CG character in the system always accelerates the convergence of FG entropy buildup. A 5-fold speedup can be attained for $\lambda = 0.25$, using the mass scaling scheme. The temperature scaling scheme provides a somewhat lower speedup, at 3.6-fold. The force addition scheme, however, provides only a moderate speedup never surpassing 1.5-fold.

The rate of CG entropy buildup is either kept or slowed down by the multiscale coupling (see the Supporting Information). Mass scaling suffered the least slowdown of CG entropy buildup for all tested λ values and could even provide a small speedup at $\lambda = 0.25$.

8.2. Conformations Sampled at the FG Level. As is apparent from Figure 1, the buildup of entropy at the FG level is faster and reaches higher values when the dynamics includes a partial CG component, both for mass and temperature scaling (colored curves, compared to black curve for $\lambda = 1$). The effect is more consistent for Martini than for IBI potentials and is virtually absent for the force addition scheme. The higher values indicate that the sampled distribution of molecular conformations is more extensive. As the method used to estimate the entropy, which assumes a multivariate Gaussian distribution of fluctuations, is only guaranteed to yield an upper bound for the actual entropy, the true entropy values may be smaller. Figure 2 clearly shows that including a CG component

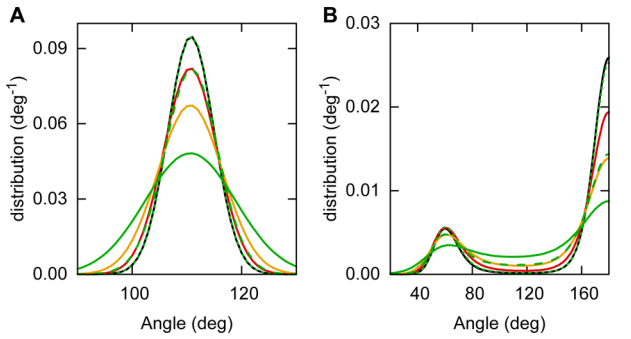


Figure 2. Conformational analysis of FG degrees of freedom. (**A**) FG angle distributions and (**B**) FG dihedral distributions, using the mass scaling scheme and the Martini CG potential. λ values go from 0.75 to 0.25 (same color code as in Figure 1, solid lines). Also depicted for comparison are the angle distributions of a pure FG system ($\lambda = 1$; full black line) and of systems at $\lambda = 0.25$ using the Martini CG potential and either the force addition scheme (dotted green line) or temperature scaling scheme (dashed green line). The departure from the pure FG behavior of the distributions obtained using the mass scaling scheme is evident. None of the other schemes displayed such deviations, even at the lowest tested λ value of 0.25. Note the extensive overlap between the pure distribution, in black, and that from the force addition scheme, in dotted green, expected due to the keeping of FG bonded forces.

in the dynamics yields a broadening of the angular distributions. This is compatible with an increase in entropy and can be interpreted as an additional disorder. We may conclude not only that in mixed FG-CG simulations the distribution in molecular configurational space becomes broader than in pure FG simulations but also that the sampling of this broader distribution is faster and more efficient. This broadening of configurational space is more evident for the mass scaling scheme, for which faithfulness to the pure FG configurations is quickly lost for $\lambda < 1$ (Figure 2).

8.3. Conformations Sampled at the CG Level. Overall, systems at intermediate λ values converge to CG entropies lower than at full FG. At $\lambda = 0$, however, CG entropy converges again to higher values: simulations with IBI potentials at $\lambda = 0$ converge to approximately the same CG entropy value as at $\lambda = 1$, whereas those with Martini potentials converge to about 3% higher values (Figure 1). Systems simulated with the force addition scheme always converge to lower entropy values for $\lambda < 1$, even for $\lambda = 0$. This is, at least in part, a consequence of FG bonded forces being kept unscaled at all λ values.

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

Of all the scheme−potential combinations, only the temperature scaling scheme using Martini preserved the entropy of the CG subsystem compared to $\lambda = 1$. All the other combinations resulted in some loss of CG entropy at λ values between 1 and 0. This decrease of CG entropy buildup is likely coupled to the also observed decrease in FG entropy buildup and to the accumulation of heat in specific d.o.f. of the CG subsystem.

8.4. System Overordering. Particular scheme/λ combinations cause the systems coupled to the IBI potentials to converge to excessively low entropies: when using either the force addition or the temperature scaling schemes, coupling at λ values closer to a full CG representation causes the system to become overly ordered. This is evident from Figures 3-A and 3-B, in which are represented the molecular c.o.m.-to-c.o.m. radial distribution functions for these two schemes using the IBI potentials. Such an increase in the ordering of the system is also evident in the CG subsystem behavior, which converges to a much lower configurational entropy buildup−translated as a much more collinear arrangement of the four CG beads (Figures 3-C and 3-D). An increase in the average mean square displacement (MSD) with decreasing λ was also observed in these cases, which is likely related to the decrease of chain entanglement brought about by the straightening of the chains at the CG level (the plots of FG and CG angle distributions, RDFs, and MSDs for all schemes and potentials are included in the Supporting Information). In at least one situation (temperature scaling scheme using IBI potentials at $\lambda = 0.25$) the strength of this effect was such that parts of the system underwent a transition to a flying ice cube as described in Section 5. It is interesting to note that while for the force addition scheme the ordering effect seems to increase with decreasing λ, for the temperature scaling scheme it only manifested itself at the lowest tested λ-value. The Martini-based systems were clearly more robust to this overordering effect. The reason why the IBI potentials allow for kinetic energy to become trapped in specific d.o.f. is perhaps related to the occurrence of softer, lower-frequency features in some of the potentials. This observation does not necessarily preclude the use of IBI potentials from multiscaling applications but did prompt the use of suitable thermostats to mitigate the effect — either local thermostats or thermostats with a stochastic component.

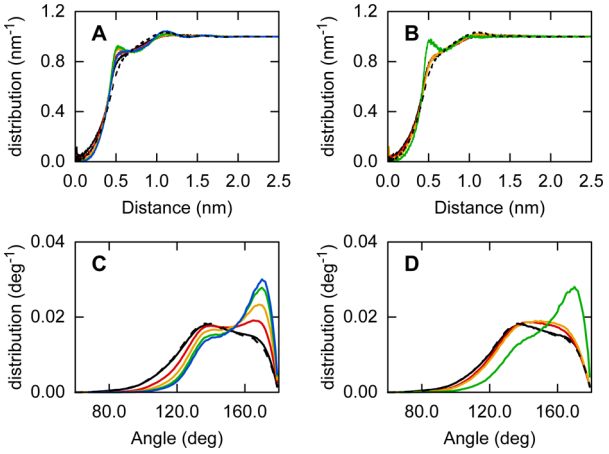


Figure 3. CG configuration analysis. (**A** and **B**) c.o.m.-to-c.o.m. radial distribution functions and (**C** and **D**) CG angle distributions using the IBI potentials for the CG subsystem. Plots depict either the force addition scheme (**A** and **C**) or the temperature scaling scheme (**B** and **D**). λ values follow the same color code as in Figure 1, and the black dashed line corresponds to a pure CG system with the IBI potentials. An ordering of the system is evident at lower λ values.

8.5. Temperature Scaling — Heat Flow from the Thermostats. The heat flow between the temperature baths and CG and FG subsystems was monitored for the temperature scaling scheme. In this scheme temperature is maintained using Berendsen thermostats[39] coupled to the CG and FG subsystems separately. The thermostat acts by scaling the velocities of all particles in a subsystem in each step by a factor λ_s (unrelated to the multiscale coupling λ), which depends on the deviation of the subsystem temperature from the required temperature. It effectively drives temperature deviations to zero through a first-order process with a time constant τ_T. If the temperature is perfectly maintained, $\lambda_s = 1$ at each step; if the subsystem tends to heat up, $\lambda_s < 1$ on average. The average values of λ_s in the temperature scaling scheme for the CG and FG subsystems at several values of λ are reported in Figure 4. It can be seen that there is a net heat flow between the bath and subsystems in the mixed CG/FG simulations using the temperature scaling thermostat. At the lower values of λ, heat flows from the thermostat out of the FG subsystem and into the CG subsystem, meaning that heat flows internally from the CG to the FG subsystems. This internal flow is expected for any CG−FG coupling because of the loss of d.o.f. — all CG movement maps to FG movement but not the other way around. In the temperature scaling scheme the effect is more visible because the FG system is kept at a lower temperature (as a comparison, mass scaling scheme exhibits λ_s values closer to 1 by 1 order of magnitude). The general trend appears to be that the net flow of heat is higher at lower λ (but vanishes at the pure CG level). In any case, the average internal heat flow is very small and drowns in the random fluctuations; temperature deviations from the bath values are only about 1% for $\lambda = 0.25$.

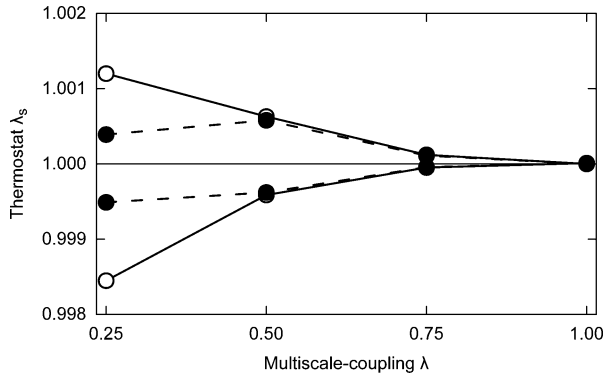


Figure 4. Thermostat scaling factor λ_s for different multiscale-coupling λ values, using the temperature scaling scheme. The CG potential is either Martini (full circles) or IBI (empty circles). Lines are drawn as a guide to the eye. The upper plots and bottom plots correspond to the CG and FG subsystem's λ_s, respectively. The estimated errors of the data points have too small a magnitude (below 10^{-6}) and were not plotted. A trend is visible where heat flow from the CG to the FG subsystem increases with lower λ values. λ_s data for the mass scaling scheme (not shown) does not display such a trend, with λ_s values being closer to 1 by an order of magnitude.

8.6. Unevenness of Heat Distribution and Stochastic Thermostats. The difference in number of d.o.f. between FG

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

and CG means that an increase in temperature of the FG subsystem need not lead to a corresponding increase in that of the CG subsystem nor to an equally well-distributed heating of the CG d.o.f. This justifies the use of separate thermostats. The unevenness in the distribution of the system heat among the subsystems and their d.o.f. further accounts for the convergence of most CG subsystems to lower entropy values and in some cases to overordered states. The effect can also be seen in systems where only the FG subsystem is directly thermostatized (like our implementation of the force addition scheme): the CG subsystem tends to a lower temperature than that of the FG subsystem. The imbalance, however, is not too drastic, as even at $\lambda = 0.25$ (using force addition with the Martini CG potential) the CG subsystem was indirectly kept at 295.4 K versus the directly thermostatized 300.0 K of the FG subsystem.

The uneven distribution of heat within the CG subsystem, which contributes to the overordering of some systems, might be mitigated by coupling the CG subsystem to a thermostat with some stochastic character. Indeed, the coupling to the SD or the DPD-ISO thermostats visibly reduces the overordering of the $\lambda = 0.25$ system simulated using the temperature scaling scheme and the IBI CG potentials (Figure 5). The speedup

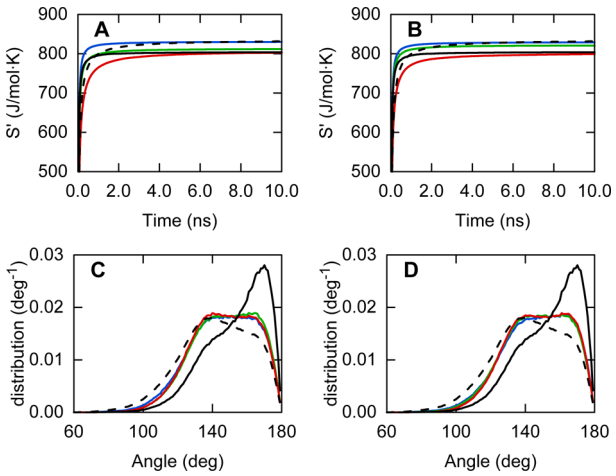


Figure 5. Entropy buildup of the FG subsystem (**A** and **B**) and CG angle distributions (**C** and **D**) for systems simulated with the temperature scaling scheme and the IBI potentials, at $\lambda = 0.25$, and with coupling to the Berendsen-SD (**A** and **C**) or Berendsen-DPD-ISO (**B** and **D**) thermostats. The blue, green, and red lines correspond to an increasing stochastic contribution of the thermostat (for Berendsen-DPD-ISO f_{iso} fractions of 0.001, 0.01, and 0.1; for Berendsen-SD f fractions of 0.999, 0.99, and 0.9; see ref 30 for details on these parameters). The lines in black, included for comparison, are the systems at $\lambda = 0.25$ (full line) and $\lambda = 1$ (dashed line) simulated using the temperature scaling scheme, IBI potentials, and the unmodified Berendsen thermostat.

performance of the multiscale scheme was kept. However, the dynamic properties of the system were disturbed, in that molecular diffusion becomes strongly dependent on the stochastic factors (see the Supporting Information). Although these factors could be tuned so that the observed diffusion matches that of the pure FG system without stochastics there is no guarantee that the dynamic details would remain the same (and they very likely would not).

9. CONCLUSIONS

Our comparative work shows that the three multiscaling schemes generate different behaviors, as does the choice of CG-subsystem potential.

Our IBI CG potential performed worse than the Martini potential in all cases, very likely due to the tendency for the accumulation of heat in specific d.o.f.

The force addition method seems to sample the FG phase space closest to that of pure FG simulations. This comes at the cost of only a relatively small speedup in entropy convergence, although this coupling method was originally proposed for this end as part of a replica-exchange scheme.[7]

Both the mass scaling and the temperature scaling schemes provide a substantial speedup of the configurational sampling at the expense of some increase in disorder compared to the pure FG behavior. Both methods are therefore suitable in hybrid FG-CG schemes to enhance sampling. Of all combinations studied, the temperature scaling method using Martini potentials preserved CG sampling the best, in that CG entropies did not decrease for $\lambda < 1$. The force addition method does not yield a significant speedup and is not recommended to enhance sampling.

Our results on hybrid schemes with constant λ (both in space and time) are encouraging enough to investigate the consequences of space- and/or time-dependent λ, thus allowing applications to systems with space and/or time-dependent resolution.

■ ASSOCIATED CONTENT

Supporting Information

A file is provided that includes the used inverted potentials as well as the structural and entropy-convergence data for all the scheme/potential combinations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: s.j.marrink@rug.nl.

Author Contributions

§These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the financial support given by The Netherlands Organization for Scientific Research (NWO) to M.N.M. (Veni Grant No. 722.013.010).

■ REFERENCES

(1) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. *WIREs Comput. Mol. Sci.* **2014**, *4*, 225−248.

(2) Saunders, M. G.; Voth, G. A. *Annu. Rev. Biophys.* **2013**, *42*, 73−93.

(3) Noid, W. *J. Chem. Phys.* **2013**, *139*, 090901.

(4) Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. *Phys. Rev. Lett.* **2005**, *95*, 218102.

(5) Praprotnik, M.; Site, L. D.; Kremer, K. *J. Comput. Chem.* **2005**, *123*, 224106.

(6) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. *Phys. Rev. Lett.* **2006**, *96*, 028105.

(7) Christen, M.; van Gunsteren, W. F. *J. Chem. Phys.* **2006**, *124*, 154106.

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398

(8) Praprotnik, M.; Kremer, K.; Site, L. D. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2007**, *75*, 017701.
(9) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. *J. Chem. Theory Comput.* **2007**, *3*, 1100−1105.
(10) Praprotnik, M.; Matysiak, S.; Delle Site, L.; Kremer, K.; Clementi, C. *J. Phys.: Condens. Matter.* **2007**, *19*, 292201.
(11) Heyden, A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 217−221.
(12) Darré, L.; Tek, A.; Baaden, M.; Pantano, S. *J. Chem. Theory Comput.* **2012**, *8*, 3880−3894.
(13) Wassenaar, T. A.; Ingólfsson, H. I.; Prieß, M.; Marrink, S. J.; Schäfer, L. V. *J. Phys. Chem. B* **2013**, *117*, 3516−3530.
(14) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. *J. Chem. Phys.* **2014**, *140*, 054114.
(15) Shen, L.; Hu, H. *J. Chem. Theory Comput.* **2014**, *10*, 2528−2536.
(16) Ayton, G. S.; Noid, W. G.; Voth, G. A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 192−198.
(17) Peter, C.; Kremer, K. *Soft Matter* **2009**, *5*, 4357−4366.
(18) Tozzini, V. *Acc. Chem. Res.* **2009**, *43*, 220−230.
(19) Kamerlin, S. C. L.; Warshel, A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10401−10411.
(20) Goga, N.; Rzepiela, A. J.; Melo, M. N.; de Vries, A. H.; Hadar, A.; Markvoort, A. J.; Nedea, S.; Berendsen, H. J. C. In *Advances in Planar Lipid Bilayers and Liposomes*; Iglic, A., Ed.; Academic Press: Burlington, MA, 2012; Vol. *15*, p 139.
(21) Li, Y.; Abberton, B. C.; Kröger, M.; Liu, W. K. *Polymers* **2013**, *5*, 751−832.
(22) Di Pasquale, N.; Gowers, R. J.; Carbone, P. *J. Comput. Chem.* **2014**, *35*, 1199−1207.
(23) van Buuren, A. R.; Marrink, S. J.; Berendsen, H. J. C. *J. Phys. Chem.* **1993**, *97*, 9206−9212.
(24) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(25) Soper, A. *Chem. Phys.* **1996**, *202*, 295−306.
(26) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(27) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *4*, 2471−2474.
(28) Berendsen, H. J. C. *Simulating the Physical World*; Cambridge University Press: Cambridge, UK, 2007; see section 19.9.3 on page 499 for the derivation of the metric tensor correction weight factor.
(29) Harvey, S. C.; Tan, R. K.-Z.; Cheatham, T. E. *J. Comput. Chem.* **1998**, *19*, 726−740.
(30) Goga, N.; Rzepiela, A. J.; de Vries, A. H.; Marrink, S. J.; Berendsen, H. J. C. *J. Chem. Theory Comput.* **2012**, *8*, 3637−3649.
(31) Peters, E. A. J. F.; Goga, N.; Berendsen, H. J. C. *J. Chem. Theory Comput.* **2014**, *10*, 4208−4220.
(32) Henderson, R. L. *Phys. Lett. A* **1974**, *49*, 197−198.
(33) Jain, S.; Garde, S.; Kumar, S. K. *Ind. Eng. Chem. Res.* **2006**, *45*, 5614−5618.
(34) Reith, D.; Pütz, M.; Müller-Plathe, F. J. *Comput. Chem.* **2003**, *24*, 1624−1636.
(35) Müller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754−769.
(36) Rühle, V.; Junghans, C.; Lukyanov, A.; Kremer, K.; Andrienko, D. *J. Chem. Theory Comput.* **2009**, *5*, 3211−3223.
(37) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(38) Hockney, R. W. *Methods Comput. Phys.* **1970**, *9*, 136−211.
(39) Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Comput. Chem.* **1984**, *81*, 3684−3690.
(40) Schlitter, J. *Chem. Phys. Lett.* **1993**, *215*, 617−621.
(41) Baron, R.; de Vries, A. H.; Hünenberger, P. H.; van Gunsteren, W. F. *J. Phys. Chem. B* **2006**, *110*, 8464−8473.

DOI: 10.1021/ct501102b
J. Chem. Theory Comput. 2015, 11, 1389−1398