The Journal of Chemical Physics

Back-mapping augmented adaptive resolution simulation

Cite as: J. Chem. Phys. **153**, 164118 (2020); https://doi.org/10.1063/5.0025728
Submitted: 18 August 2020 . Accepted: 09 October 2020 . Published Online: 28 October 2020

S. Thaler, M. Praprotnik, and J. Zavadlav

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Back-mapping augmented adaptive resolution simulation

Cite as: J. Chem. Phys. 153, 164118 (2020); doi: 10.1063/5.0025728
Submitted: 18 August 2020 • Accepted: 9 October 2020 •
Published Online: 28 October 2020

S. Thaler,[1] M. Praprotnik,[2,a)] and J. Zavadlav[1,b)]

AFFILIATIONS

[1] Professorship of Multiscale Modeling of Fluid Materials, Department of Mechanical Engineering, Technical University of Munich, Munich, Germany
[2] Laboratory for Molecular Modeling, National Institute of Chemistry, SI-1001 Ljubljana, Slovenia

[a)] **Also at:** Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, SI-1000 Ljubljana, Slovenia.
[b)] **Author to whom correspondence should be addressed:** julija.zavadlav@tum.de

ABSTRACT

Concurrent multiscale techniques such as Adaptive Resolution Scheme (AdResS) can offer ample computational advantages over conventional atomistic (AT) molecular dynamics simulations. However, they typically rely on aphysical hybrid regions to maintain numerical stability when high-resolution degrees of freedom (DOFs) are randomly re-inserted at the resolution interface. We propose an Energy Minimized AT (DOF) Insertion (EMATI) method that uses an informed rather than random AT DOF insertion to tackle the root cause of the issue, i.e., overlapping AT potentials. EMATI enables us to directly couple AT and coarse-grained resolutions without any modifications of the interaction potentials. We exemplify AdResS-EMATI in a system of liquid butane and show that it yields improved structural and thermodynamic properties at the interface compared to competing AdResS approaches. Furthermore, our approach extends the applicability of the AdResS without a hybrid region to systems for which force capping is inadequate.

I. INTRODUCTION

Biomolecular processes are challenging for computational modeling as they involve a vast span of time and length scales as the macroscopic properties of interest emerge from a molecular origin. While coarse-grained (CG) representations can reach larger spatiotemporal scales,[1,2] these models lack the accuracy and detail of atomistic (AT) models. Several multiscale approaches aim at resolving these conflicting objectives including back-mapping methods,[3–7] resolution replica exchange methods,[8,9] and concurrent multiscale simulations. The latter approach incorporates an intriguing idea of a computational magnifying glass: preserving atomistic accuracy and detail around a region of interest while reducing the remainder of the system to its essential degrees of freedom (DOFs).[10] Prototypical applications are processes where atomistic details are of interest only in a localized region, e.g., binding processes[11] and interactions of antimicrobial peptides with lipid membranes.[12] The coupling of AT and CG representations can be utilized either with constant resolution methods[13–16] or adaptive resolution methods.[10,17–26]

In the Adaptive Resolution Scheme[17] (AdResS), the simulation domain is separated into an AT and a CG region. Particles can diffuse freely between both regions, changing their DOFs on the fly. To allow for a smooth change in the resolution of the transitioning particles, a hybrid (HY) region is introduced at the interface of the AT and CG regions. However, the inclusion of the HY region is computationally demanding as it requires the computation of forces from both AT and CG potentials.[27,28] Furthermore, structural properties deviate in the HY region even if they match in the AT and CG regions.[29] First attempts to overcome these drawbacks were made in a recent paper by Krekeler *et al.*,[28] where the AdResS was employed in the limiting case of no HY region and sizable computational speed-ups over the standard AdResS were reported.

However, direct coupling of different resolutions results in interface difficulties that are in the standard AdResS alleviated by the HY region. A central hindrance is due to potentially overlapping AT particles when CG sites migrate from the CG into the AT region.[28] The current implementations of the AdResS in GROMACS[28,30] and Espresso++[31] insert AT DOFs randomly. Due to this random scheme, atoms are sometimes inserted unphysically close to existing atoms in the AT region, resulting in fatally high forces from steric repulsion that ultimately make the simulation numerically unstable.

To avoid these fatally large forces, Krekeler *et al.*[28] simply capped forces above a specified threshold, even though the authors noted that a method yielding proper AT DOFs might be necessary in certain cases. Up to now, the method was only applied to small and/or rather spherical solvent molecules.[21,28,32] Spherical molecules do not tend to yield severe overlaps as a spherical CG potential can be a good approximation to the AT molecule.[25] By contrast, non-spherical AT molecules can extend significantly beyond the van der Waals (VdW) volume enforced by the CG potential as we discuss in this paper. For these systems, random placement of DOFs at the interface can be detrimental to the AdResS even with the HY region.[25] For example, in an AdResS simulation of alkane systems,[25] the Lennard-Jones (LJ) potential needed to be substituted by a soft-core potential in the HY region that gradually blends back to the original LJ potential toward the AT region. Such modifications for the sake of avoiding fatally large forces, however, alter the AT force field and hence also the properties close to the AT-CG interface.

In this paper, we propose the Energy Minimized AT (DOF) Insertion (EMATI) method that avoids fatally overlapping potentials by using an informed rather than random AT DOF insertion. We, therefore, tackle the root cause of the numerical instability instead of inserting DOFs randomly and *ad hoc* mitigating the consequences of occasional overlaps. To solve the problem of finding valid AT DOFs based on the center of mass (COM) and the surrounding chemical environment, we transfer methods from the closely related back-mapping multiscale approach[3–6] to the AdResS. We demonstrate that direct coupling of AT and CG resolutions without the HY region with AdResS-EMATI eliminates numerical instability in a system of liquid butane without requiring force capping or AT potential modifications. By contrast, simulations with force capping become numerically unstable because occasionally more energy is introduced into the system at the interface than can be dissipated by the thermostat for common friction values. Additionally, we showcase improved interface properties with our method compared to both the standard AdResS with the HY region (subsequently referred to as AdResS) and AdResS without the HY region using force capping[28] (subsequently referred to as AdResS-FC).

II. METHODS

A. Adaptive resolution simulation

The AdResS[17] divides the simulation domain in an AT, a CG, and a HY region. The total force $\mathbf{F}_\alpha$ acting on a molecule α is

$$\mathbf{F}_\alpha = \sum_{\beta\neq\alpha} w(\mathbf{R}_\alpha)w(\mathbf{R}_\beta)\mathbf{F}^{AT}_{\alpha\beta} + \sum_{\beta\neq\alpha}[1 - w(\mathbf{R}_\alpha)w(\mathbf{R}_\beta)]\mathbf{F}^{CG}_{\alpha\beta} + \mathbf{F}^{TD}_\alpha,$$
$$\text{with} \quad w(\mathbf{R}) = \begin{cases} 1 \text{ if } \mathbf{R} \in \text{AT} \\ 0 \text{ if } \mathbf{R} \in \text{CG} \\ 0 < w(\mathbf{R}) < 1 \text{ if } \mathbf{R} \in \text{HY}, \end{cases} \tag{1}$$

where $\mathbf{R}$ is a molecule's COM, $\mathbf{F}^{AT}_{\alpha\beta}$ and $\mathbf{F}^{CG}_{\alpha\beta}$ are forces acting between molecules α and β via the AT and CG force fields, respectively, and $w(\mathbf{R})$ is a smooth resolution weighting function. $\mathbf{F}^{TD}$ is a thermodynamic force, which compensates differences in chemical potentials between both resolutions[22,29,33] and is typically applied in a close neighborhood of the resolution interface.

Shrinking the size of the HY region to 0 (Fig. 1) conceptually transforms $w(\mathbf{R})$ into the Heaviside step function,[28] reducing its purpose to a switching function that sorts molecules into AT or CG resolution. In this case, Eq. (1) implies that two molecules with the same resolution interact via the force field of the respective resolution, whereas molecules with different resolutions interact via the CG force field. Such a coupling definition is reminiscent of constant resolution multiscale methods,[13–16] whose common feature is direct interaction of molecules at different resolutions. In particular, the virtual sites approach[14,34] models the AT-CG interaction via the unaltered CG force field, equivalently to the AdResS without the HY region, albeit not allowing molecules to change their resolution. Furthermore, without the HY region, the AdResS[17] becomes a Hamiltonian method.[24] As already mentioned, omitting the HY region can cause the simulation to become numerically unstable unless the AT DOFs are inserted in a proper way. Section II B describes the EMATI scheme.

B. EMATI scheme

The aim of the presented EMATI method is to insert AT DOFs at sensible locations such that no fatally large forces occur in an AdResS simulation without the HY region. For each molecule entering the AT region, EMATI therefore needs to propose valid AT DOFs based on neighboring atoms in the AT region and given a

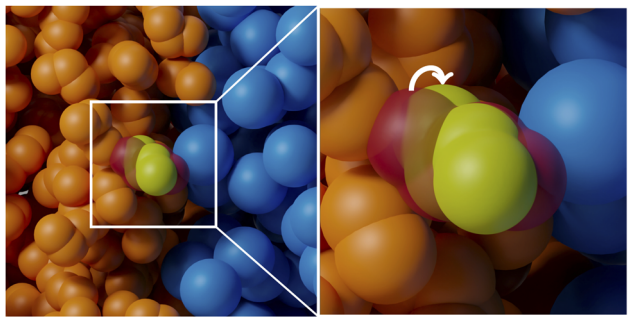

FIG. 1. Visualization of the interface of the AdResS for liquid butane without a HY region. The AT region resolves butane molecules atomistically (orange), while the CG region only resolves the COM of each molecule (blue). Random DOF insertion may yield potential overlaps (red molecule). EMATI proposes DOFs based on surrounding AT molecules avoiding severe overlaps (yellow molecule).

fixed COM position defined by the CG site. A well-known algorithm for on-the-fly insertion of molecules into dense fluids for open system simulations is USHER.[35] A generalization of USHER[36] achieves molecule insertion at prescribed potential energy values by simultaneously moving the COM and rigidly rotating the molecule using a steepest descent iterator. However, this approach relies on adjusting the COM of molecules to be inserted to find an appropriate insertion position. USHER is therefore not applicable to the AdResS, where the COM of the molecule is fixed. Furthermore, in its original formulation,[36] it does not generalize to non-rigid molecules. The insertion task in the AdResS resembles much more closely the central problem of the back-mapping multiscale approach,[3,4] where the AT detail needs to be back-inserted into a given skeleton of CG sites. Two of the most common ingredients of back-mapping algorithms are random initial insertion of atoms[5,6] and energy minimization to avoid overlapping AT potentials,[3,4,6,7] which serve as the main components of the EMATI method.

Our AdResS-EMATI approach tracks when a molecule has migrated from the CG into the AT region and applies EMATI (Fig. 2) to each of those CG sites. The first step of EMATI is to retrieve all neighboring AT atoms within a cutoff r_{cut} around the CG site position $\mathbf{R}$. This retrieval of neighbors is only required once as all particle positions are fixed, except for the molecule whose AT DOFs we are inserting (subsequently referred to as the central molecule). Note that r_{cut} can be chosen smaller than the cutoff of the AT potential to increase computational efficiency because a small number of nearest atoms dominate repulsive forces.

Algorithm 1: Energy Minimized AT DOF Insertion (EMATI)

input : A CG site position $\mathbf{R}$ and initial AT positions $\{\mathbf{r}_i^0\}_{i=1}^N$
output: Valid AT positions $\{\mathbf{r}_i^{\text{final}}\}_{i=1}^N$
$\{\mathbf{r}_i^{\text{Nei}}\}_{i=1}^{N_{\text{Nei}}}$ = gather_AT_neighbors($\mathbf{R}$, r_{cut});
for resets = 1 **to** n_{resets} **do**
  while $\sigma > \sigma_{\text{target}}$ **do**
    $\{\mathbf{F}_i\}_{i=1}^N$ =sum_forces($\{\mathbf{r}_i^{\text{Nei}}\}_{i=1}^{N_{\text{Nei}}}$, $\{\mathbf{r}_i^k\}_{i=1}^N$);
    $F^{\text{COM}} = \|\sum_{i=1}^N \mathbf{F}_i\|$;
    $F^{\max} = \max(\{\|\mathbf{F}_i\|\}_{i=1}^N)$;
    σ = compute_SD(F^{COM}, n_{window}) ;
    foreach $\mathbf{r}_i^k$ *in* $\{\mathbf{r}_i^k\}_{i=1}^N$ **do** $\mathbf{r}_i^{k+1} = \mathbf{r}_i^k + \frac{\alpha \mathbf{F}_i}{m_i \max(1, F^{\max}/F^{\text{thresh}})}$;
    $\mathbf{R}^{\text{new}}$ = compute_COM($\{\mathbf{r}_i^k\}_{i=1}^N$) ;
    foreach $\mathbf{r}_i^k$ *in* $\{\mathbf{r}_i^k\}_{i=1}^N$ **do** $\mathbf{r}_i^{k+1} = \mathbf{r}_i^{k+1} - \mathbf{R}^{\text{new}} + \mathbf{R}$;
    $k = k + 1$;
  end
  if $F^{\text{COM}} < F^{\text{target}}$ **then**
    $\{\mathbf{r}_i^{\text{final}}\}_{i=1}^N = \{\mathbf{r}_i^k\}_{i=1}^N$;
    break;
  else if resets = n_{resets} **then**
    $\{\mathbf{r}_i^{\text{final}}\}_{i=1}^N = \{\mathbf{r}_i^{\text{cur\_best}}\}_{i=1}^N$;
  else
    if $F^{\text{COM}} < F^{\text{cur\_best}}$ **then**
      $\{\mathbf{r}_i^{\text{cur\_best}}\}_{i=1}^N = \{\mathbf{r}_i^k\}_{i=1}^N$;
      $F^{\text{cur\_best}} = F^{\text{COM}}$;
    end
    $\{\mathbf{r}_i^0\}_{i=1}^N$ = random_reset($\{\mathbf{r}_i^k\}_{i=1}^N$) ;
  end
end

FIG. 2. Energy minimized AT DOF insertion (EMATI) algorithm.

The inner gradient-based constrained potential energy minimization loop is the core of EMATI. It serves to find a local minimum of the potential energy for the central molecule while satisfying the CG site position constraint. As the starting configuration for the energy minimization $\{\mathbf{r}_i^0\}_{i=1}^N$, where N is the number of AT particles per molecule, we choose the pseudo-random insertion provided by the AdResS implementation. The pseudo-randomness emerges from AdResS implementations that do not delete AT DOFs upon leaving the AT region. Inside the CG region, AT particles travel along with the CG site and simply re-appear at their current positions upon re-entering the AT region. We additionally implemented a truly random initial insertion but found no measurable effect on results. We opt for a steepest descent energy minimization scheme with step size α. The potential energy gradient is computed from forces acting on the central molecule from AT inter- and intra-molecular interactions. Displacing the current atom positions of the central molecule $\{\mathbf{r}_i^k\}_{i=1}^N$ along the steepest descent direction yields the configuration of the next iteration step,

$$\mathbf{r}_i^{k+1} = \mathbf{r}_i^k + \frac{\alpha \mathbf{F}_i}{m_i \max(1, F^{\max}/F^{\text{thresh}})}, \tag{2}$$

where m_i is the mass of the particle i and $\mathbf{F}_i$ is the force on particle i exerted by the neighboring atoms. To avoid overshooting local minima, we re-scale all forces if the magnitude of the maximum force $F^{\max} = \max(\{\|\mathbf{F}_i\|\}_{i=1}^N)$ exceeds a prescribed threshold force $F^{\max} > F^{\text{thresh}}$. This gradient re-scaling guarantees a constant maximum atomic displacement per iteration step. Updating AT positions according to Eq. (2) changes the COM of the central molecule to $\mathbf{R}^{\text{new}}$. To fulfill the "mapping condition,"[4] i.e., that the central molecule COM coincides with the CG site position, we move the central molecule back to the original $\mathbf{R}$ in each iteration, exactly fulfilling this constraint.

We define the convergence criterion of the constrained energy minimization based on the variance of the gradient: A local minimum is obtained when the standard deviation of the magnitude of the COM force $F^{\text{COM}} = \|\sum_{i=1}^N \mathbf{F}_i\|$ of the last n_{window} steps is smaller than the target σ_{target}. This variance-based convergence criterion is more suitable than directly dictating a maximum F^{COM} because it allows detection of local minima where further energy minimization would not yield a significantly better configuration, thus saving computational effort. A convergence criterion similar to Ref. 35 based on the potential could also be formulated. However, we opt for the above-mentioned criterion based on forces to avoid the additional computation of potential values. A convergence criterion based on the displacement of the central molecule might be a reasonable alternative.

The outer resetting loop checks if the obtained local minimum is acceptable, i.e., if $F^{\text{COM}} < F^{\text{target}}$. Otherwise, the molecule is reset randomly to yield new initial atom positions $\{\mathbf{r}_i^0\}_{i=1}^N$ to search for a better local minimum. We save $\{\mathbf{r}_i\}_{i=1}^N$ that yielded the smallest F^{COM} to continue the simulation with the best obtained configuration in case none of the obtained configurations yields a $F^{\text{COM}} < F^{\text{target}}$. In our simulations, we found this resetting scheme to be necessary for numerical stability as it avoids being stuck in unacceptable local minima, a known phenomenon in USHER[35,36] and back-mapping problems.[6] Theoretically, a series of n_{resets} unfortunate random resets that all lead to local minima with fatally

large forces would be possible. However, increasing n_{resets} significantly reduces the probability of this rare event. Note that this situation did not occur in our implementation of EMATI, even for simulations of 10 ns length with more than 300 000 EMATI executions.

The presented method shows similarities to the rejection criterion in Monte Carlo Simulations in the sense that the unlikely configuration from random insertion is rejected and the higher likelihood configuration output from EMATI is accepted.[28]

C. Simulation setup

We choose liquid butane as an exemplary solvent to demonstrate the effectiveness of AdResS-EMATI. We use the GROMOS53A5[37] force field with flexible bond lengths to represent AT butane. The CG potential was derived via Iterative Boltzmann Inversion (IBI) including pressure correction[38] with the STOCK coarse-graining kit.[39] We performed a 10 ns AT reference simulation in a 5 × 5 × 5 nm^3 box to compute the target AT radial distribution function (RDF) and pressure. The obtained CG potential is shown in Fig. 3. Both AT and CG potentials are cut off at 1.4 nm.

We perform all simulations in the NVT ensemble in an orthorhombic simulation box with periodic boundary conditions using the Espresso++ 2.0.2.[31] package. We use a velocity Verlet time integration scheme with a 2 fs time step in accordance with the GROMOS[37] force field and a Langevin thermostat to maintain a target temperature of 323 K. We choose a friction coefficient of $\gamma = 1/\text{ps}$ unless stated otherwise.

For multiscale simulations, we use the box size of 20 × 5 × 5 nm^3 containing 3174 molecules, which corresponds to a density of 612.7 kg/m^3. For AdResS-EMATI and AdResS-FC, the simulation box is split along the x axis with an AT region of length 10 nm in the center and two connected CG regions (due to periodicity) of 5 nm (Fig. 1). For AdResS simulations, we choose an AT region width of 7.2 nm and a HY region length of 1.4 nm such that the explicit HY region coincides with the interface region of AdResS-EMATI, where AT molecules are influenced by AT and CG force fields. All simulations are run for 10 ns to generate the data.

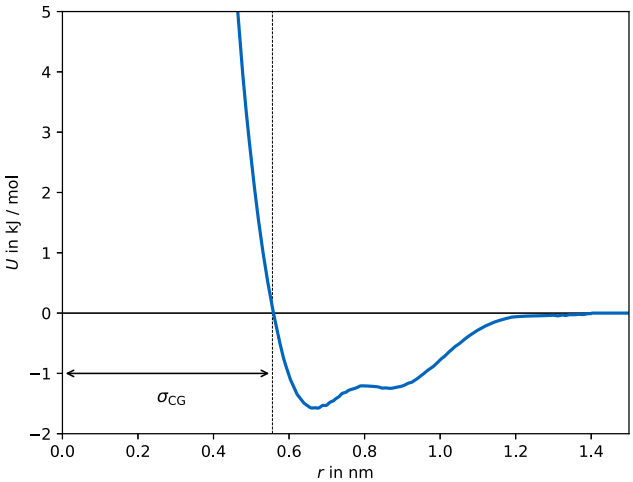

FIG. 3. CG potential of liquid butane obtained by iterative Boltzmann inversion. We defined σ_{CG} analogous to the LJ potential.

$\mathbf{F}^{\text{TD}}$ acts in a close neighborhood of the resolution interface and guarantees a uniform particle density distribution by construction. It is obtained iteratively before the production run $\mathbf{F}^{TD}_{i+1}(x) = \mathbf{F}^{TD}_{i}(x) - C\nabla\rho_i(x)$, where C is a convergence-driven, tunable constant.[29,40] We performed 30 iterations of length 1 ns each to reach a converged $\mathbf{F}^{\text{TD}}$ for AdResS-EMATI and AdResS-FC, while 50 iterations were necessary for the AdResS. An iteration constant $C = 2.2 \cdot 10^{-3}$ (kJ m^3)/(mol kg) was selected.

To maintain numerical stability with the AdResS and AdResS-FC, we cap forces component-wise at $F^{\text{cap}} = 5000$ kJ/(mol nm). For the AdResS-EMATI, we track the migration of molecules across the resolution regions via the Heaviside switching function $w(\mathbf{R})$. The EMATI scheme is triggered for all molecules whose w switches from 0 to 1. The scheme is applied after the velocity Verlet position update that caused the resolution change but before the force re-computation, where valid AT DOFs are necessary. We implement EMATI with the following parameters: $\alpha = 28.125$ fs^2, $r_{\text{cut}} = r_{\text{cut,AT}}/2 = 0.7$ nm, $F^{\text{thresh}} = 2000$ kJ/(mol nm), $\sigma_{\text{target}} = 5$ kJ/(mol nm), $n_{\text{window}} = 5$, $F^{\text{target}} = 1000$ kJ/(mol nm), and $n_{\text{resets}} = 10$. The random resetting function rotates the central molecule rigidly around a randomly drawn axis by a random angle between 45° and 135°. This range of angles ensures a large rotation to avoid staying inside the same insufficient local minimum.

III. RESULTS AND DISCUSSION

A. AdResS-EMATI

We demonstrate that AdResS-EMATI fulfills three central requirements: AT and CG regions are in equilibrium, it reproduces structural properties of a reference AT simulation, and the resolution interface does not act as an artificial diffusion barrier. Figure 4 visualizes the normalized density profile (NDP) of AdResS-EMATI with $\mathbf{F}^{\text{TD}}$ as a function of the distance from the AdResS center d. For the employed $\mathbf{F}^{\text{TD}}$, see Fig. 8. A maximum error of less than 2.5% in the NDP confirms good convergence of $\mathbf{F}^{\text{TD}}$. This minor density deviation together with a close to homogeneous temperature profile (discussed below) shows that AT and CG regions are in equilibrium.

We analyze the quality of reproducing structural properties by computing the COM–COM RDF (Fig. 5). The RDF in the AT region matches the reference full AT RDF perfectly, and the very well fit of the RDF in the CG region confirms convergence of the IBI. Note that both AT and CG RDFs are computed based on structural data from their whole respective domain, without ignoring particles in a neighborhood around the resolution interface. Shrinking the HY region to 0, therefore, significantly increases the size of the domain usable for structural analyses. For a more detailed discussion of the structural quality in the interface region, see Fig. 9.

Figure 6 demonstrates the absence of an artificial diffusion barrier in AdResS-EMATI. Particles in the AT region up to 1 nm left of

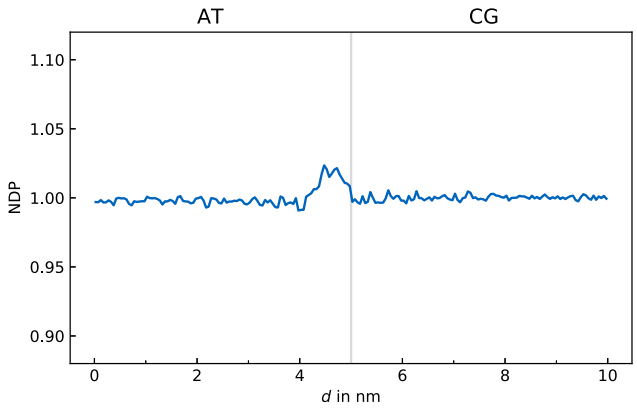

FIG. 4. Normalized density profile (NDP) of AdResS-EMATI with $\mathbf{F}^{TD}$. The gray line visualizes the AT-CG interface.

the interface are marked at $t = 0$, and the probability density functions (PDFs) of marked particles over time are visualized. The larger PDF tails in the CG region reflect the higher diffusion constant in the CG region due to smoother dynamics from a lack of fluctuating forces that are missing AT DOFs.[41] The PDFs clearly show undisturbed Brownian motion, confirming that EMATI does not hinder diffusion.

B. Comparison to AdResS-FC and AdResS

AdResS-FC does not yield numerically stable simulations for $\gamma \leq 5$ because, occasionally, more energy is introduced into the system at the interface from capped overlap forces than the thermostat can dissipate. This monotonically rising temperature makes the simulation numerically unstable. We therefore show results for $\gamma = 6$ in order to compare to AdResS-FC, even though such high values can affect the dynamics and viscosity of the system.[42,43] The common friction values for butane multiscale simulations in the literature are, e.g., $\gamma = 0.1$/ps[44] or $\gamma = 1$/ps.[14] Note that we additionally tested AdResS-FC with the MARTINI CG potential and $\gamma = 1$, resulting in the same type of fatal temperature rise. By contrast,

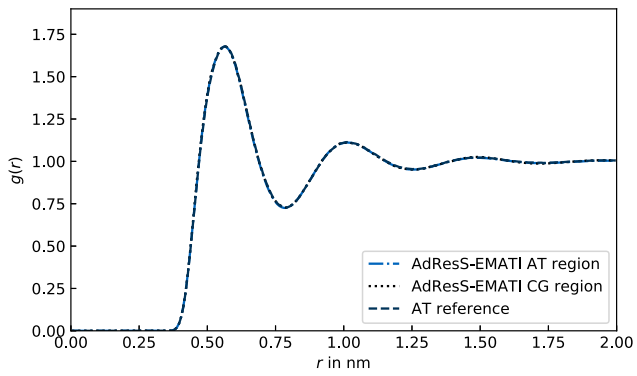

FIG. 5. COM–COM RDFs of AdResS-EMATI in the AT and CG region compared to a reference full AT simulation.

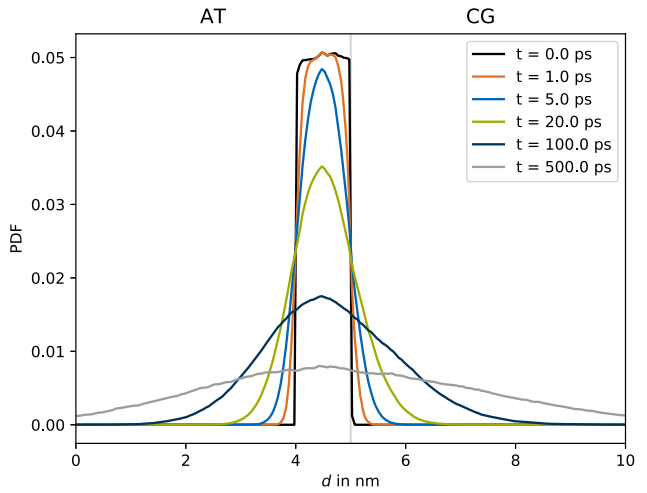

FIG. 6. Particle diffusion across the resolution interface for AdResS-EMATI. The AT-CG interface is visualized by a gray line. Particles in the AT region next to the interface are marked at $t = 0$, and the PDFs of marked particles over time are visualized.

AdResS-EMATI yields numerically stable simulations with both CG potentials even for $\gamma = 0.1$/ps, showing no evidence of these fatally rising temperatures.

A constant temperature profile is as important as a constant density profile for AT and CG regions to be in equilibrium.[22,29,33,45] We compare the normalized temperature profile (NTP) of AdResS-EMATI and AdResS-FC in Fig. 7 for $\gamma = 6$ to allow a comparison on equal terms. For the AdResS-EMATI, $\gamma = 6$ simulation, we reuse $\mathbf{F}^{TD}$ derived for $\gamma = 1$. Despite the large friction coefficient of $\gamma = 6$, AdResS-FC yields a peak temperature deviation of 10.8%, while AdResS-EMATI only yields a maximum deviation of 0.6% for the same friction value. For $\gamma = 1$, AdResS-EMATI yields a temperature error of 3.6%, which is on the same order as the density deviation of 2.4%.

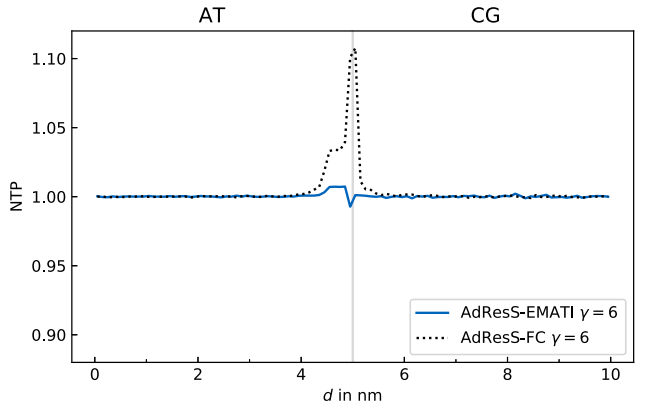

FIG. 7. Normalized temperature profile (NTP) of AdResS-EMATI and AdResS-FC. The temperature is computed based on CG DOFs in the CG region and AT DOFs in the AT region. The gray line visualizes the AT-CG interface.

As the temperature deviation is primarily controlled by the Langevin thermostat, a dedicated γ for the region around the resolution interface might improve AdResS-FC: The effect of a necessarily larger γ would be limited to this region, while the AT and CG regions could be simulated with a more desirable, smaller γ. Nonetheless, the above conclusions from our simulations with a constant γ still hold in the sense that AdResS-FC always requires a much larger γ in the interface region than AdResS-EMATI for the same maximum deviation threshold. This larger than desired impact of the thermostat might eventually alter structural and/or thermodynamic properties around the interface.

Figure 8 displays $\mathbf{F}^{\mathrm{TD}}$ and the NDP without $\mathbf{F}^{\mathrm{TD}}$ for AdResS-EMATI and AdResS-FC. NDP and $\mathbf{F}^{\mathrm{TD}}$ are very similar for both methods except for a very narrow region around the interface, hinting at a strictly local impact of EMATI and force capping. The most striking difference in NDP is a larger density well directly at the interface for AdResS-FC, presumingly due to large forces from overlaps (capped at F^{cap}) causing both overlapping molecules to quickly diffuse away from the interface. The same phenomenon emerges in $\mathbf{F}^{\mathrm{TD}}$ as it needs to compensate for this density well, resulting in larger force extrema close to the interface.

While achieving a constant NTP is not problematic in the AdResS, the smoothing inside the HY region comes at the cost of structural deviations in this region. Figure 9 compares the COM–COM RDF of AdResS-EMATI at the region that coincides with the HY region of the AdResS. The agreement with the AT reference RDF is very good for AdResS-EMATI (and for AdResS-FC, not shown). By contrast, the AdResS yields an RDF in the HY region that significantly deviates from the AT reference.

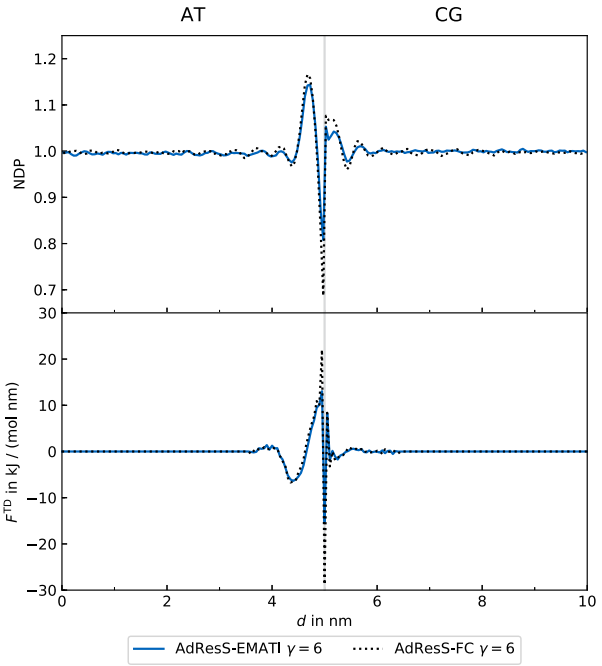

FIG. 8. Normalized density profile (NDP) and $\mathbf{F}^{\mathrm{TD}}$ of AdResS-EMATI and AdResS-FC. The gray line visualizes the AT-CG interface.

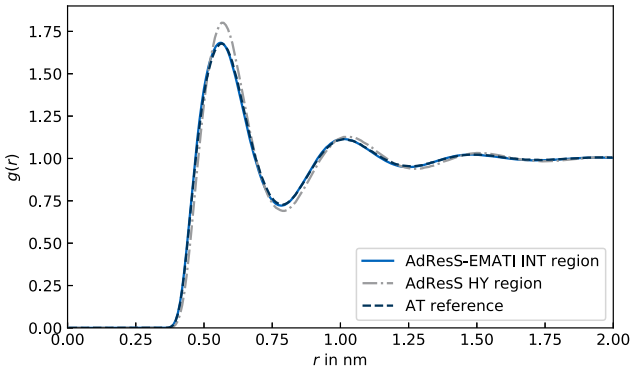

FIG. 9. COM–COM RDFs of AdResS-EMATI in the interface (INT) region compared to the AdResS in the HY region and to the reference full AT simulation.

Our results demonstrate that AdResS-EMATI yields well matching thermodynamic as well as structural properties at the interface. Matching additional properties beyond density and temperature at the interface increases numerical accuracy in AT and CG regions by reducing coupling artifacts.[29] For example, matching the COM–COM RDF in the interface implies that the PDF in the AT region matches the PDF of a full AT simulation at least up to second order,[22,29] while even third-order accuracy has been shown empirically.[22] This characteristic of reduced AT-CG coupling artifacts makes AdResS-EMATI a prime choice in a computational magnifying glass setting where the CG region serves to deliver a coarse but informative representation of the system.

C. Estimation of overlap severity

We propose a coarse measure based on the solvent geometry and both the AT and CG force fields to *a priori* estimate the possible severity of overlapping potentials at the interface. This measure might be helpful in identifying systems for which augmenting the AdResS with EMATI (or some other similar back-mapping method) is necessary. Note that we neglect electrostatic forces in this discussion as steric repulsion forces from the LJ potential are dominant for small atom distances.

The superposition of the VdW volumes of all likely occurring AT configurations that correspond to a given COM position yields a sphere of radius r_{AT} that determines the maximum extent of inserted AT molecules [Fig. 10(a)]. We approximate r_{AT} as a sum of the VdW radius of the outermost AT atom $\sigma_{\mathrm{AT}}/2$ and its distance from the COM position in the equilibrium configuration[46] r_{COM}, i.e., $r_{\mathrm{AT}} \approx r_{\mathrm{COM}} + \sigma_{\mathrm{AT}}/2$. The difference between the radius of this AT sphere and the VdW radius of the CG potential ($r_\delta = r_{\mathrm{AT}} - \sigma_{\mathrm{CG}}/2$) determines the likelihood and severity of overlaps for random insertion, where we estimate σ_{CG} analogously to the LJ potential (Fig. 3).

We define a measure for overlap severity ν by considering the worst possible scenario [Fig. 10(b)]. Suppose two neighboring CG

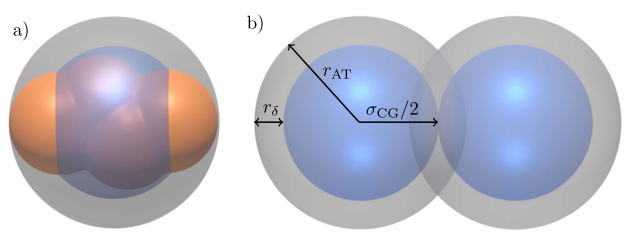

FIG. 10. Sketch of the severity of overlap estimation. Panel (a) visualizes the origin of the sphere of maximum AT extent for butane. The VdW volume of a sample butane molecule in orange significantly extends beyond the VdW volume of the CG potential in blue. Superposition of all possible AT configurations yields the gray sphere of maximum AT extent. Panel (b) sketches the configuration of the worst case overlap estimation with two molecules being σ_{CG} apart. This configuration yields overlapping AT spheres shaded dark gray. If both AT molecules happen to point exactly toward each other, their most outward VdW spheres overlap by $2r_\delta$.

sites are separated by a distance σ_{CG}. Coincidentally, the AT DOFs are inserted such that the outermost AT atoms of the two CG sites are facing each other, i.e., they are at the minimal mutual distance equal to $r = \sigma_{CG} - 2r_{COM} = \sigma_{AT} - 2r_\delta$. The overlap severity v is

$$v = \frac{F^{AT}|_{r=\sigma_{AT}-2r_\delta}}{F^{AT}|_{r=\sigma_{AT}}}, \tag{3}$$

where $F^{AT} = \|\mathbf{F}^{AT}\|$ is the force magnitude computed from the AT potential of the molecule's outermost AT atom.

In the case of SPC water, where v is not excessively large ($r_\delta = 0.026$ nm, $v = 17.1$), even this worst case insertion scenario does not create fatally large forces. Therefore, force capping is not necessary for SPC water in the AdResS without the HY region,[47] which we confirmed with test simulations. Larger, non-spherical molecules are more challenging for the AdResS, given that a spherical CG potential is not a good approximation[25] and r_{AT} increases with the maximum distance of the outermost atoms from the COM position r_{COM}. Thus, alkane chains, including butane,[44] require a method to handle overlapping AT potentials, even with the HY region.[25] The issue is even worse without the HY region with v of butane ($r_\delta = 0.10$ nm, $v = 5.4 \cdot 10^4$) being more than three orders of magnitude larger than in the case of SPC water.

Any back-mapping method that augments the AdResS implicitly relies on the assumption that the COM of the central molecule is sensible, given the surrounding AT neighbors. Examples that clearly violate this assumption are the recent applications of the AdResS without the HY region modeling CG molecules as ideal gas particles:[21,32] Ideal gas particles in the CG region do not enforce a minimum distance between CG molecules ($\sigma_{CG} = 0$); hence, the COM of the central molecule can be arbitrarily close to neighboring AT atoms. Consequently, there does not always exist an AT configuration of the central molecule that is consistent with the given COM and simultaneously yields non-fatal forces, necessitating force capping, even for water.[32]

The main drawback of EMATI is its reliance on energy minimization. Energy minimization is prone to yield over-stabilized structures in back-mapping problems such that additional molecular dynamics steps are often required to obtain structures at a target state point.[4] However, we did not experience large over-stabilization effects in our simulations. If, however, application of EMATI would result in unacceptably large over-stabilization, increasing σ_{target} should shift the potential energy distribution toward larger energies counteracting this effect.

Energy minimization steps also increase computational effort, generating a computational overhead over AdResS-FC. The extra computational effort is proportional to the number of CG sites migrating into the AT region, hence scaling with the resolution interface area. In our simulations, EMATI is applied approximately once every 17 time steps. Computation per migrated CG site, i.e., per EMATI execution, is limited to a small hemisphere of radius r_{cut} around the central molecule and is mainly controlled by the EMATI parameters r_{cut}, σ_{target}, F^{target}, and n_{resets}. On the other hand, AdResS-EMATI reduces the overhead due to the exclusion of the HY region. The related speed-up scales with the volume of the HY region and depends on the employed force fields. For example, Ref. 28 reported a speed-up of 1.4 for a system of two micelles in water. We refer the reader to Refs. 27 and 28 for a detailed discussion of the computational implications of the HY region. In our numerical experiments, the overhead of EMATI, with the parameters given in Sec. II C, was approximately the same as the overhead from the HY region. For this proof of concept study, we neither optimized EMATI nor its parameters for numerical efficiency. As a test, we also changed EMATI's parameters toward numerical efficiency (e.g., reducing r_{cut}) and found a significant reduction in overhead, hence outperforming the AdResS. The numerical efficiency could also be improved by considering more advanced methods, e.g., the Fast Inertial Relaxation Engine (FIRE).[48]

IV. CONCLUSION

In this work, we proposed EMATI, an interface DOF insertion method for the AdResS without the need of a HY region sandwiched in between the AT and CG regions. AdResS-EMATI is conceptually similar to the method introduced in Ref. 28. However, it largely extends the applicability of the direct AT/CG coupling. In particular, it enables the direct resolution coupling to systems with non-spherical molecules, e.g., butane, for which simple force capping does not suffice. Numerical stability in AdResS-EMATI is achieved without requiring force capping or changing the AT potential at the interface. AdResS-EMATI directly tackles the root cause of the numerical instability issue, i.e., overlapping AT potentials, by inserting AT DOFs based on the minimized interacting energy with surrounding atoms. We demonstrated the applicability of our method in a system of liquid butane, for which AdResS-FC results in a fatal temperature rise for common thermostat friction values. We further showcased reduced temperature artifacts over AdResS-FC while also removing the structural discrepancies observed for the standard AdResS in the HY region.

We chose the IBI CG potential to showcase excellent structural properties in the interface region (Fig. 9). However, there

are many CG potentials that would work with AdResS-EMATI as well, e.g., a potential derived from the multiscale coarse-graining method[49–51] would be a reasonable alternative. We additionally implemented the MARTINI[52] CG force field and found the same numerical stability preserving properties of AdResS-EMATI, even though the RDFs of the MARTINI and AT force fields do not match. AdResS-EMATI is in principle applicable to polar and apolar solvents and arbitrary CG force fields, as long as the CG potential enforces a sufficiently large VdW volume such that an acceptable AT molecule configuration exists. It could also be applied to coarse-grained models of molecules with several CG particles, i.e., coarse-grained models of polymers. In this case, the calculation of the forces used for the EMATI scheme needs to also include the bonded interactions (bonds, angles, and dihedrals) between the atoms belonging to different CG sites in the AT region of the same molecule.

EMATI might also prove valuable to the standard AdResS with the HY region, e.g., for macromolecular systems,[25] as an alternative to the soft-core potential substitution method. Augmentation of the AdResS without the HY region by a back-mapping method represents one step toward its applicability as a computational magnifying glass by improving numerical stability and reducing artifacts from overlaps in the interface. To improve the method further, replacing the energy minimization in EMATI by a more advanced back-mapping scheme[53–57] might be a next step worth investigating.

ACKNOWLEDGMENTS

M.P. acknowledges financial support through Grant No. P1-0002 from the Slovenian Research Agency.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems," J. Chem. Phys. **139**, 090901 (2013).
[2]H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole, and S. J. Marrink, "The power of coarse graining in biomolecular simulations," Wiley Interdiscip. Rev. Comput. Mol. Sci. **4**, 225–248 (2014).
[3]A. P. Heath, L. E. Kavraki, and C. Clementi, "From coarse-grain to all-atom: Toward multiscale analysis of protein landscapes," Proteins Struct. Funct. Genet. **68**, 646–661 (2007).
[4]C. Peter and K. Kremer, "Multiscale simulation of soft matter systems—From the atomistic to the coarse-grained level and back," Soft Matter **5**, 4357–4366 (2009).
[5]A. J. Rzepiela, L. V. Schäfer, N. Goga, H. Jelger Risselada, A. H. De Vries, and S. J. Marrink, "Software news and update reconstruction of atomistic details from coarse-grained structures," J. Comput. Chem. **31**, 1333–1343 (2010).
[6]T. A. Wassenaar, K. Pluhackova, R. A. Böckmann, S. J. Marrink, and D. P. Tieleman, "Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models," J. Chem. Theory Comput. **10**, 676–690 (2014).
[7]M. Shimizu and S. Takada, "Reconstruction of atomistic structures from coarse-grained models for protein-DNA complexes," J. Chem. Theory Comput. **14**, 1682–1694 (2018).
[8]E. Lyman, F. M. Ytreberg, and D. M. Zuckerman, "Resolution exchange simulation," Phys. Rev. Lett. **96**, 028105 (2006).
[9]P. Liu and G. A. Voth, "Smart resolution replica exchange: An efficient algorithm for exploring complex energy landscapes," J. Chem. Phys. **126**, 045106 (2007).
[10]M. Praprotnik, L. Delle Site, and K. Kremer, "Multiscale simulation of soft matter: From scale bridging to adaptive resolution," Annu. Rev. Phys. Chem. **59**, 545–571 (2008).
[11]Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, and D. E. Shaw, "How does a drug molecule find its target binding site?," J. Am. Chem. Soc. **133**, 9181–9183 (2011).
[12]Y. Wang, T. Zhao, D. Wei, E. Strandberg, A. S. Ulrich, and J. P. Ulmschneider, "How reliable are molecular dynamics simulations of membrane active antimicrobial peptides?," Biochim. Biophys. Acta **1838**, 2280–2288 (2014).
[13]Q. Shi, S. Izvekov, and G. A. Voth, "Mixed atomistic and coarse-grained molecular dynamics: Simulation of a membrane-bound ion channel," J. Phys. Chem. B **110**, 15045–15048 (2006).
[14]A. J. Rzepiela, M. Louhivuori, C. Peter, and S. J. Marrink, "Hybrid simulations: Combining atomistic and coarse-grained force fields using virtual sites," Phys. Chem. Chem. Phys. **13**, 10437–10448 (2011).
[15]P. Sokkar, S. M. Choi, and Y. M. Rhee, "Simple method for simulating the mixture of atomistic and coarse-grained molecular systems," J. Chem. Theory Comput. **9**, 3728–3739 (2013).
[16]A. B. Kuhn, S. M. Gopal, and L. V. Schäfer, "On using atomistic solvent layers in hybrid all-atom/coarse-grained molecular dynamics simulations," J. Chem. Theory Comput. **11**, 4460–4472 (2015).
[17]M. Praprotnik, L. Delle Site, and K. Kremer, "Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly," J. Chem. Phys. **123**, 224106 (2005); arXiv:0510223 [cond-mat].
[18]J. Zavadlav, M. N. Melo, S. J. Marrink, and M. Praprotnik, "Adaptive resolution simulation of an atomistic protein in martini water," J. Chem. Phys. **140**, 054114 (2014).
[19]J. Zavadlav, S. J. Marrink, and M. Praprotnik, "Multiscale simulation of protein hydration using the SWINGER dynamical clustering algorithm," J. Chem. Theory Comput. **14**, 1754–1761 (2018).
[20]J. Zavadlav, J. Sablić, R. Podgornik, and M. Praprotnik, "Open-boundary molecular dynamics of a DNA molecule in a hybrid explicit/implicit salt solution," Biophys. J. **114**, 2352–2362 (2018).
[21]J. Whittaker and L. Delle Site, "Investigation of the hydration shell of a membrane in an open system molecular dynamics simulation," Phys. Rev. Res. **1**, 033099 (2019).
[22]H. Wang, C. Hartmann, C. Schütte, and L. Delle Site, "Grand-canonical-like molecular-dynamics simulations by using an adaptive-resolution technique," Phys. Rev. X **3**, 011018 (2013).
[23]A. Agarwal, H. Wang, C. Schütte, and L. Delle Site, "Chemical potential of liquids and mixtures via adaptive resolution simulation," J. Chem. Phys. **141**, 034102 (2014); arXiv:1311.6982.
[24]R. Potestio, S. Fritsch, P. Español, R. Delgado-Buscalioni, K. Kremer, R. Everaers, and D. Donadio, "Hamiltonian adaptive resolution simulation for molecular liquids," Phys. Rev. Lett. **110**, 108301 (2013).
[25]J. H. Peters, R. Klein, and L. Delle Site, "Simulation of macromolecular liquids with the adaptive resolution molecular dynamics technique," Phys. Rev. E **94**, 023309 (2016).
[26]A. Chaimovich, C. Peter, and K. Kremer, "Relative resolution: A hybrid formalism for fluid mixtures," J. Chem. Phys. **143**, 243107 (2015); arXiv:1903.04755.
[27]C. Junghans, A. Agarwal, and L. Delle Site, "Computational efficiency and Amdahl's law for the adaptive resolution simulation technique," Comput. Phys. Commun. **215**, 20–25 (2017).
[28]C. Krekeler, A. Agarwal, C. Junghans, M. Praprotnik, and L. Delle Site, "Adaptive resolution molecular dynamics technique: Down to the essential," J. Chem. Phys. **149**, 024104 (2018); arXiv:1806.09870.
[29]H. Wang, C. Schütte, and L. Delle Site, "Adaptive resolution simulation (AdResS): A smooth thermodynamic and structural transition from atomistic to coarse grained resolution and *vice versa* in a grand canonical fashion," J. Chem. Theory Comput. **8**, 2878–2887 (2012).

[30]M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," SoftwareX **1-2**, 19–25 (2015).
[31]H. V. Guzman, N. Tretyakov, H. Kobayashi, A. C. Fogarty, K. Kreis, J. Krajniak, C. Junghans, K. Kremer, and T. Stuehn, "ESPResSo++ 2.0: Advanced methods for multiscale molecular simulation," Comput. Phys. Commun. **238**, 66–76 (2019); arXiv:1806.10841.
[32]L. Delle Site, C. Krekeler, J. Whittaker, A. Agarwal, R. Klein, and F. Höfling, "Molecular dynamics of open systems: Construction of a mean-field particle reservoir," Adv. Theory Simul. **2**, 1900014 (2019).
[33]S. Poblete, M. Praprotnik, K. Kremer, and L. Delle Site, "Coupling different levels of resolution in molecular simulations," J. Chem. Phys. **132**, 114101 (2010).
[34]Y. Liu, A. H. De Vries, J. Barnoud, W. Pezeshkian, J. Melcr, and S. J. Marrink, "Dual resolution membrane simulations using virtual sites," J. Phys. Chem. B **124**, 3944–3953 (2020).
[35]R. Delgado-Buscalioni and P. V. Coveney, "USHER: An algorithm for particle insertion in dense fluids," J. Chem. Phys. **119**, 978–987 (2003); arXiv:0303366 [cond-mat].
[36]G. De Fabritiis, R. Delgado-Buscalioni, and P. V. Coveney, "Energy controlled insertion of polar molecules in dense fluids," J. Chem. Phys. **121**, 12139 (2004).
[37]C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6," J. Comput. Chem. **25**, 1656–1676 (2004).
[38]D. Reith, M. Pütz, and F. Müller-Plathe, "Deriving effective mesoscale potentials from atomistic simulations," J. Comput. Chem. **24**, 1624–1636 (2003).
[39]S. Bevc, C. Junghans, and M. Praprotnik, "STOCK: Structure mapper and online coarse-graining kit for molecular simulations," J. Comput. Chem. **36**, 467–477 (2015).
[40]S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, and K. Kremer, "Adaptive resolution molecular dynamics simulation through coupling to an internal particle reservoir," Phys. Rev. Lett. **108**, 170602 (2012); arXiv:1112.3151.
[41]S. Matysiak, C. Clementi, M. Praprotnik, K. Kremer, and L. Delle Site, "Modeling diffusive dynamics in adaptive resolution simulation of liquid water," J. Chem. Phys. **128**, 024503 (2008).
[42]K. Kremer and G. S. Grest, "Dynamics of entangled linear polymer melts: A molecular-dynamics simulation," J. Chem. Phys. **92**, 5057–5086 (1990).
[43]C. Junghans, M. Praprotnik, and K. Kremer, "Transport properties controlled by a thermostat: An extended dissipative particle dynamics thermostat," Soft Matter **4**, 156–161 (2008).
[44]J. Zavadlav, M. N. Melo, A. V. Cunha, A. H. de Vries, S. J. Marrink, and M. Praprotnik, "Adaptive resolution simulation of MARTINI solvents," J. Chem. Theory Comput. **10**, 2591–2598 (2014).
[45]M. Praprotnik, K. Kremer, and L. Delle Site, "Adaptive molecular resolution via a continuous change of the phase space dimensionality," Phys. Rev. E **75**, 017701 (2007).
[46]Allowing for flexible bonds and angles will increase r_{AT} in practice.
[47]B. Duenweg, J. Castagna, S. Chiacchiera, H. Kobayashi, and C. Krekeler (2018). "Meso- and multi-scale modelling E-CAM modules II," Zenodo. https://doi.org/10.5281/zenodo.1210075 (2018).
[48]E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, "Structural relaxation made simple," Phys. Rev. Lett. **97**, 170201 (2006).
[49]S. Izvekov and G. A. Voth, "A multiscale coarse-graining method for biomolecular systems," J. Phys. Chem. B **109**, 2469–2473 (2005).
[50]W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, "The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models," J. Chem. Phys. **128**, 244114 (2008).
[51]W. G. Noid, P. Liu, Y. Wang, J.-W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth, "The multiscale coarse-graining method. II. Numerical implementation for coarse-grained molecular models," J. Chem. Phys. **128**, 244115 (2008).
[52]S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. De Vries, "The MARTINI force field: Coarse grained model for biomolecular simulations," J. Phys. Chem. B **111**, 7812–7824 (2007).
[53]L.-J. Chen, H.-J. Qian, Z.-Y. Lu, Z.-S. Li, and C.-C. Sun, "An automatic coarse-graining and fine-graining simulation method: Application on polyethylene," J. Phys. Chem. B **110**, 24093–24100 (2006).
[54]J. Krajniak, Z. Zhang, S. Pandiyan, E. Nies, and G. Samaey, "Reverse mapping method for complex polymer systems," J. Comput. Chem. **39**, 648–664 (2018).
[55]J. Peng, C. Yuan, R. Ma, and Z. Zhang, "Backmapping from multiresolution coarse-grained models to atomic structures of large biomolecules by restrained molecular dynamics simulations using Bayesian inference," J. Chem. Theory Comput. **15**, 3344–3353 (2019).
[56]G. Zhang, A. Chazirakis, V. A. Harmandaris, T. Stuehn, K. C. Daoulas, and K. Kremer, "Hierarchical modelling of polystyrene melts: From soft blobs to atomistic resolution," Soft Matter **15**, 289–302 (2019).
[57]W. Li, C. Burkhart, P. Polińska, V. Harmandaris, and M. Doxastakis, "Backmapping coarse-grained macromolecules: An efficient and versatile machine learning approach," J. Chem. Phys. **153**, 041101 (2020).