Title

Accurate and general treatment of electrostatic interaction in Hamiltonian adaptive resolution simulations

Journal

European Physical Journal: Special Topics, 225(8-9)

ISSN

1951-6355

Authors

Heidari, M
Cortes-Huerto, R
Donadio, D
et al.

Publication Date

2016-10-01

DOI

10.1140/epjst/e2016-60151-6

Peer reviewed

Eur. Phys. J. Special Topics

DOI: 10.1140/epjst/e2016-60151-6

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Regular Article

Accurate and general treatment of electrostatic interaction in Hamiltonian adaptive resolution simulations

M. Heidari[1], R. Cortes-Huerto[1], D. Donadio[2,1], and R. Potestio[1,a]

[1] Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany
[2] Department of Chemistry, University of California Davis, One Shields Ave. Davis, CA, 95616, USA

Received 30 April 2016 / Received in final form 16 June 2016
Published online 18 July 2016

Abstract. In adaptive resolution simulations the same system is concurrently modeled with different resolution in different subdomains of the simulation box, thereby enabling an accurate description in a small but relevant region, while the rest is treated with a computationally parsimonious model. In this framework, electrostatic interaction, whose accurate treatment is a crucial aspect in the realistic modeling of soft matter and biological systems, represents a particularly acute problem due to the intrinsic long-range nature of Coulomb potential. In the present work we propose and validate the usage of a short-range modification of Coulomb potential, the Damped shifted force (DSF) model, in the context of the Hamiltonian adaptive resolution simulation (`H-AdResS`) scheme. This approach, which is here validated on bulk water, ensures a reliable reproduction of the structural and dynamical properties of the liquid, and enables a seamless embedding in the `H-AdResS` framework. The resulting dual-resolution setup is implemented in the `LAMMPS` simulation package, and its customized version employed in the present work is made publicly available.

1 Introduction

The definition of soft matter encompasses a broad spectrum of different systems, from liquids composed by single atoms or simple polymer molecules [1–4] to large and complex biomolecular assemblies [5–16]. The structural and dynamical properties of this ample variety of systems spans an equivalently wide range of length and time scales, the interplay of which gives rise to a wealth of different properties.

This multi-scale nature, however, poses substantial challenges to modeling and simulation. *In silico* experiments are limited by the size of the system and the duration that can be reached by the simulation. It is often the case that the process of interest takes place in a typical time interval that is not within the reach of present day computers. An alternative to overcome such limitation is to employ a simpler

[a] e-mail: potestio@mpip-mainz.mpg.de

description of the system, where the level of detail is reduced, for example replacing a fully atomistic description and representing a group of different atoms as a single interaction site. These coarse-grained (CG) models [17–22] have provided an extraordinary tool to made mesoscale systems accessible by simulations over time scales that would not be viable through an atomistic description. Nonetheless, there are several circumstances in which it is not possible to investigate a certain system or process by means of a CG model, for example when chemically specific interactions play a substantial role.

In recent years several strategies have been developed to find a compromise between model accuracy and computational efficiency. A prominent example of these approaches are adaptive, dual-resolution simulation schemes [23–33]. These methods identify a specific, typically small region of the system that necessitates modeling at full detail, e.g. atomistic (AT). This subregion is indeed described with the high-resolution model required, while in the remainder of the system a simpler, effective CG representation is employed. In modeling liquid systems, a crucial feature is that the two regions at different resolutions are connected by an open boundary, which allows the diffusion of molecules (e.g. solvent particles) from one subdomain to the other. The instantaneous location of a molecule specifies the interactions with the neighboring molecules, thus allowing the resolution of a molecule to change on the fly.

Adaptive resolution methods thus enable the simulation of a system whose description is not bound to be the same everywhere, rather it is more accurate where strictly necessary, and simpler and computationally more efficient in the rest. Two advantages can be envisaged: on the one hand, there is an obvious gain in terms of simulation time, due to the reduced number of degrees of freedom and the simpler interactions that are employed in the low-resolution subdomain. On the other hand, these setups can be employed to characterize the physical properties of the system by systematically changing the size of the high-resolution region, so to effectively probe the locality of physical phenomena and finite size effects. This strategy enables the controllable decoupling between the internal degrees of freedom of a chosen subregion of the system from the rest, yet without modifying the thermodynamical equilibrium in the high resolution domain [34,35].

Among the methods that have implemented this strategy, a notable place is occupied by the adaptive resolution simulation (`AdResS`) [23–26] and the Hamiltonian adaptive resolution simulation (`H-AdResS`) [29,30,32,33] schemes. In these setups, schematically represented in Fig. 1, the resolution of a molecule is determined by the value of a function, usually dubbed *switching function* λ, that is equal to 1 in the high-resolution or atomistic (AT) subregion, 0 in the low-resolution or coarse-grained (CG) subregion, and smoothly interpolates between these values in an interface region, dubbed hybrid region (HY). When a molecule is in the AT or CG domains it is treated as fully atomistic or fully coarse-grained, respectively; in the HY region the interactions are obtained by interpolation of the AT and CG ones.

The `AdResS` and `H-AdResS` methods have been validated on various systems and in different contexts, and have been shown to effectively and efficiently provide an accurate description of the system in the AT region at a lower computational cost with respect to an equivalent simulation employing the high-resolution model everywhere. However, a fundamental problem affects these schemes and limits the range of systems that are amenable by them, namely the treatment of long-range interactions, such as electrostatics. Devising an efficient implementation of electrostatics in computer simulations of systems with periodic boundary conditions is still an open problem, due to the long-range nature of Coulomb potential. Nonetheless, well-established techniques are presently available to deal with these interactions in a variety of different physical contexts. In particular, the Ewald summation (ES) [36] method and its subsequent

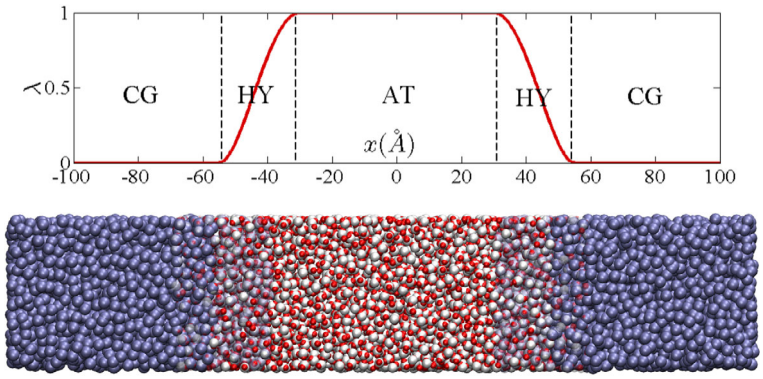

Fig. 1. Setup of a Hamiltonian Adaptive Resolution Simulation. The periodic box is partitioned into three different regions, namely: Coarse-grained (CG), Hybrid (HY), and Atomistic (AT). Upper panel: the switching function λ takes values between 0 (CG) and 1 (AT), thus defining the resolution of a molecule (here water). Lower panel: simulation snapshot explicitly showing the various subdomains.

variations provide a theoretically sound procedure to decompose the interaction in a local term, that is treated as a conventional short-range potential, and a non-local, long-range term that is efficiently computed in Fourier space.

This second term is the critical one, as it is intrinsically incompatible with the space-dependent decomposition of interactions that characterizes dual-resolution simulations. In fact, the latter approaches rely on the short-range nature of interaction potentials in order to provide a local definition of molecular resolution. In the presence of a non-local term, which requires the simultaneous knowledge of the position of each particle to compute the forces, this position-based separation of the interactions is not anymore possible.

The strategy that has been employed so far to circumvent this problem is to locally approximate the Coulomb interaction by the Reaction Field (RF) [37–39] potential. The assumption underlying this approach is that beyond a fixed (short) cutoff distance the effective long range electrostatic contribution is equivalent to that of a uniform and homogeneous dielectric medium. The standard form of the potential is thus replaced by a mean-field function, thereby reducing the interaction to a short-range one. The RF method has been successfully employed in dual-resolution setups [35, 40], however it suffers from two substantial limitations. The first one is that the assumption of a uniform and homogeneous medium beyond the cutoff distance is not always valid: this would be the case for heterogeneous interfaces, for example, polar molecules in proximity of a metal surface, or large biomolecules (protein, DNA filament) in solution [41]. The second limitation is that the parametrization of the RF potential necessitates the previous knowledge of the relative dielectric constant of the medium, which is not always available *a priori* and would then have to be computed in an independent simulation. It may also be undesirable to introduce as a parameter of the model a quantity that is indeed an emergent property of the system. Additionally, it has been shown that an accurate modeling of the system under examination sometimes necessitates a specific parametrization not only of the dielectric constant, but also of the underlying force field [39].

Here we approach the problem of electrostatic interaction in dual-resolution, adaptive simulations making use of an alternative formulation of Coulomb potential, namely the damped shifted force (DSF) potential [42, 43] method. This strategy allows us to rephrase the electrostatic interaction in terms of finite-ranged, two-body

analytical potentials, as in the case of the RF. Albeit the computational cost of RF and DSF would be identical if the same cutoff range is employed, the DSF allows us to circumvent the limitations intrinsic in both the ES [44–46] and RF methods.

It is generally the case, when making use of adaptive resolution strategies, that one saves simulation time at the cost of performing an accurate (and computationally cumbersome) parametrization of the setup, in terms of CG model, approach-specific parameters, and, as in the case just discussed, a very limited treatment of electrostatic interaction in the high-resolution model. Here we make use of elegant solutions to these problems that have been already developed and successfully validated, and can therefore be seamlessly implemented and employed. In the case of DSF, for example, the use of this technique would prove to be even advantageous over Ewald summations for some cases of interest to the computational biochemistry community. Furthermore, we present and discuss the DSF approach to the `H-AdResS` scheme in the comprehensive framework of the implementation of the latter in the `LAMMPS` [47] simulation package. The goal of this work is thus twofold: to present and validate the DSF approach in the context of `H-AdResS` simulations, and to present the implementation of these methods in the `LAMMPS` package, together with some advanced features whose practical implementation is here discussed in detail for the first time.

The manuscript is organized as follows. In Sect. 2 we discuss in detail the `H-AdResS` scheme, with a particular focus on the implementation features; then we briefly report the formulation of the RF method and the DSF method. In Sect. 3 we list and define the quantities that have been employed to validate the efficacy and accuracy of the proposed model. In Sect. 4 we report the details of the setups and the simulations, while in Sect. 5 we present the results of our study. The conclusions and perspectives are discussed in Sect. 6. The Appendix, Sect. 6, provides a list of technical details on the LAMMPS implementation of the `H-AdResS` method and its usage.

2 Methods

In this Section we review the different computational techniques employed in the present work. The first part is devoted to the `H-AdResS` scheme. Most of the fundamental and conceptual aspects of this method have been thoroughly discussed in previous publications [29,30,32,33]. The focus is here given to the computational aspects of the scheme and, in particular, of the specific implementation in the `LAMMPS` [47] simulation package.[1]

Subsequently, we summarize two of the three methods here used to treat electrostatic interaction, specifically the Reaction Field and the DSF methods. The Particle-Particle Particle-Mesh (PPPM) [48] Ewald summation method, not discussed in detail here, is taken as the golden standard, against which the results of the other two strategies are compared.

2.1 `H-AdResS`

In the `H-AdResS` scheme, the description of the interactions within a system of particles is given in terms of a global Hamiltonian function $\mathcal{H}$, which has the following

[1] The `H-AdResS` method is now a standard feature of the `LAMMPS` simulation package and can be cloned from the Git repository `https://github.com/lammps/lammps`. Alternatively, the `LAMMPS` version featuring `H-AdResS` can be freely downloaded from the address `http://www2.mpip-mainz.mpg.de/~potestio/software.php`.

form:

$$H = \mathcal{K} + V^{int} + \sum_{\alpha} \left\{ \lambda_\alpha V_\alpha^{AT} + (1-\lambda_\alpha) V_\alpha^{CG} \right\}. \tag{1}$$

The term $\mathcal{K}$ is the atomistic kinetic energy, and V^{int} consists of all the intramolecular bonded interactions (e.g. bond stretching). The resolution of particle α is specified by the transition function $\lambda_\alpha = \lambda(\mathbf{R}_\alpha)$, which is computed on the center-of-mass coordinates $\mathbf{R}_\alpha$ of the molecule.

The value of the switching function is determined by the sizes d_{at} and d_{hy} of the AT and HY regions, respectively, and of the specific geometry of the AT region. If the latter is defined as a slab of the simulation box, for example as it is represented in Fig. 1, d_{at} will correspond to the width of the atomistic subdomain; if a spherical geometry is employed, d_{at} will correspond to the AT region diameter. In all cases, the value of d_{hy} indicates the width or thickness of the HY layer embedding the AT region.

In the present work we employ a rectangular simulation box, and the AT region is a slab located in the middle of it. The resolution of a molecule is then determined through the following piece-wise λ function:

$$\lambda(x) = \begin{cases} 1 & |x| \leq d_{at}/2 \\ \cos^2\left(\frac{\pi(x-d_{at}/2)}{2d_{hy}}\right) & \frac{d_{at}}{2} < |x| \leq \frac{d_{at}}{2} + d_{hy} \\ 0 & |x| > d_{at} + d_{hy}. \end{cases} \tag{2}$$

The mid point of the simulation box is set in the origin of the coordinate system. As it is shown in Fig. 1, in all simulations the width of the AT and HY region is set to $d_{at} = 60\,\text{Å}$ and $d_{hy} = 25\,\text{Å}$, respectively.

A molecule interacts with its neighboring particles through coarse-grained V^{CG} and atomistic V^{AT} potentials. The functional form of these potentials is arbitrary, as well as the order of the interaction (two-body, three-body...) as long as the extension of the interaction is finite and short-ranged. For simplicity, in the following we shall restrict the discussion to the most common case of pairwise interactions. In the Hamiltonian of Eq. (1) the non-bonded potential energy contribution of each molecule α is given by a weighted sum of two terms V_α^{CG} and V_α^{AT}, defined as:

$$\begin{aligned} V_\alpha^{AT} &\equiv \frac{1}{2} \sum_{\beta,\beta\neq\alpha}^{N} \sum_{ij} V^{AT}(|\mathbf{r}_{\alpha i} - \mathbf{r}_{\beta j}|) \\ V_\alpha^{CG} &\equiv \frac{1}{2} \sum_{\beta,\beta\neq\alpha}^{N} V^{CG}(|\mathbf{R}_\alpha - \mathbf{R}_\beta|). \end{aligned} \tag{3}$$

The AT and CG terms of each molecule are weighted by λ_α or $(1-\lambda_\alpha)$, respectively. As the total non-bonded potential is given by the sum of this linear combination over all molecules, V_α^{AT} and V_α^{CG} contain a factor $1/2$ to account for the double counting. The corresponding force acting on atom i of molecule α is given by:

$$\begin{aligned} \mathbf{F}_{\alpha i} = \mathbf{F}_{\alpha i}^{int} &+ \sum_{\beta,\beta\neq\alpha} \left\{ \frac{\lambda_\alpha + \lambda_\beta}{2} \mathbf{F}_{\alpha i|\beta}^{AT} + \left(1 - \frac{\lambda_\alpha + \lambda_\beta}{2}\right) \mathbf{F}_{\alpha i|\beta}^{CG} \right\} \\ &- \left[V_\alpha^{AT} - V_\alpha^{CG}\right] \nabla_{\alpha i} \lambda_\alpha. \end{aligned} \tag{4}$$

The first term, $\mathbf{F}_{\alpha i}^{int}$, is due to the interactions with atoms in the same molecule, and is not subject to resolution-dependent reweighting; the second term is the sum, over all other molecules β in the interaction range, of the pairwise atomistic and coarse-grained forces, weighted by the average resolutions of the two molecules. This term is antisymmetric under molecule exchange, and satisfies Newton's Third Law by construction. The last term emerges as a consequence of the non-uniformity of space in the dual-resolution simulation setup, that is, the fact that different interactions are present in different parts of the system. Because of this, translational invariance is locally broken, and a force emerges in the hybrid region (where $\nabla\lambda \neq 0$) and acts on the molecules pushing them in one of the two subdomains, depending on the sign of the prefactor $(V_\alpha^{AT} - V_\alpha^{CG})$.

The terms V_α^{AT} and V_α^{CG} contain the different potentials acting on the molecules, and in most cases they represent an atomistic potential that takes into account chemical specificity, and and effective, coarse-grained potential, which acts on collective degrees of freedom of the molecules (e.g. the center of mass). The CG potentials are normally potentials of mean force that enclose entropic contributions and are parametrized over specific thermodynamic properties, e.g. the radial distribution function (RDF), at a specific state point. These CG interactions and the corresponding reference system follow different equations of state, and, once coupled together via an open boundary as in the case of `H-AdResS`, they exchange particles to balance the differences in equilibrium pressure and chemical potential. As it has been already thoroughly investigated [29,30,33], models that, for the same temperature and density, attain different pressure, will determine in the dual-resolution setup a non-homogeneous density, as the region where the pressure is higher will relax by pushing molecules in the other region. Furthermore, in the `H-AdResS` setup the aforementioned *drift force* term $\mathbf{F}_{\alpha i}^{dr} = -\left[V_\alpha^{AT} - V_\alpha^{CG}\right]\nabla_{\alpha i}\lambda_\alpha$ contributes to determine an imbalance in the pressure across the HY region, as it pushes molecules in the subdomain where Helmholtz free energy is locally lower [29,30].

To overcome these effects and enforce a uniform density profile, it is possible to introduce a new term in the Hamiltonian:

$$H_\Delta = H - \sum_{\alpha=1}^{N} \Delta H(\lambda(\mathbf{R}_\alpha)). \tag{5}$$

This term acts separately on each molecule in the system and plays two roles: it removes, on average, the drift force, and enforces a uniform density profile by imposing, in each subdomain, the pressure at which each model has, separately, the correct density. In the following we discuss the computational techniques employed to parametrize the term $\Delta H(\lambda)$.

2.1.1 Compensation of the drift force (pressure route)

In order to remove, on average, the effect of the drift force, the compensation term $\Delta H(\lambda)$ has to satisfy the relation:

$$\left.\frac{d\Delta H(\lambda)}{d\lambda}\right|_{\lambda=\lambda_\alpha} = \left\langle\left[V_\alpha^{AT} - V_\alpha^{CG}\right]\right\rangle_{\mathbf{R}_\alpha}. \tag{6}$$

If this is the case, the total drift force resulting from the modified Hamiltonian reads:

$$\hat{\mathbf{F}}_\alpha^{dr} = \left(V_\alpha^{AT} - V_\alpha^{CG} - \left.\frac{d\Delta H(\lambda)}{d\lambda}\right|_{\lambda=\lambda_\alpha}\right)\nabla\lambda\left(\mathbf{R}_\alpha\right) \tag{7}$$

and by construction $\langle \hat{\mathbf{F}}_\alpha^{dr} \rangle = 0$. It has been shown [29,33] that by compensating the drift force the hydrostatic pressure is uniform across the whole simulation domain, while in each of the two regions the densities may differ, as they are the equilibrium ones at that pressure according to the equation of state of each model. To compute the appropriate value of the compensation function it is possible to perform a Kirkwood thermodynamic integration (KTI) [49], since, as it has been demonstrated [29,33], the potential $\Delta H(\lambda)$ can be approximated in a mean field fashion by the Helmholtz free energy difference between a system with hybrid Hamiltonian at resolution λ and the reference (CG) system with $\lambda = 0$. However, this procedure requires a free energy calculation just to parametrize the compensation term, and its accuracy can be limited when the correlations within the hybrid region are too strong.

A more effective strategy is to compute and balance the drift force locally and parametrize the compensation *on the fly* within an iterative scheme [33]. The HY region is discretized according to the resolution λ in a number N_b of bins of width $\delta\lambda = 1/N_b$. For each molecule α in a given bin $i = \texttt{floor}[\lambda(\mathbf{R}_\alpha)/\delta\lambda]$ of the HY region, the V_α^{CG} and V_α^{AT} terms are computed and accumulated in separated local variables V_i^{CG} and V_i^{AT}, respectively; also variables N_i^{CG} and N_i^{AT} are defined to keep track of the number of molecules present in the bin. These computations are performed simultaneously and in the same routine where the CG and AT forces are calculated (See 6.3 and 6.4).

This procedure is carried out for all molecules within the HY region and continues for a time interval of duration Δt. At the end of the n-th interval the average AT and CG potential terms are computed, defined as $\bar{V}^R[i,n] = V_i^R/N_i^R$, where the index $R \in \{AT, CG\}$ specifies the resolution. The variables V_i^R, N_i^R are emptied and the averaging procedure continues. The average values calculated at the end of the n-th interval Δt are employed to compute the running average $\mathcal{V}_{i,n}^R$ of the terms, that is:

$$\mathcal{V}_{i,n+1}^R = \frac{n\ \mathcal{V}_{i,n}^R + \bar{V}^R[i,n]}{n+1} \tag{8}$$

where n is initialized at 0 and $\mathcal{V}_{i,0}^R = 0$.

For $n > 0$, the running average terms are employed to compute separately the different components of the force needed to compensate the drift force. Specifically, at time t such that $t_0 + n\Delta t < t \leq t_0 + (n+1)\Delta t$ a molecule located in bin i of the HY region will experience the following compensation forces:

$$\mathbf{F}_{\alpha,i}^R = s\ \mathcal{V}_{i,n}^R \nabla\lambda\,(\mathbf{R}_\alpha) \tag{9}$$

where $s = +1$ if $R = AT$ and $s = -1$ if $R = CG$. At each time step this force is spread to the atoms of the molecule proportionally to the relative mass of the atom over the mass of the molecule (see Eq. (23)).

The running average update continues until the compensation forces have converged to a stable value in each bin i; after this point, the update is interrupted and the resulting compensation is given by a time-independent, resolution-based force field that can be integrated to compute the corresponding contribution $\Delta H(\lambda)$ to the total energy of the system.

2.1.2 Compensation of the density imbalance (density route)

The application of the compensation of the drift force in the HY region enforces a uniform pressure profile across the whole system. However, due to the different

equations of state of the AT and CG models, at equilibrium a density gradient between the two main subdomains will arise. When one needs a uniform density profile in the simulation box, it is necessary to modify the compensation term ΔH in order to establish, in each subregion, the appropriate pressure at which the different models attain the same density.

This correction can be obtained in an iterative scheme dubbed *thermodynamic force* calculation [28], consisting in successively applying to the molecules in the HY region a force proportional to the density gradient:

$$\mathbf{F}^{th}_{n+1} = \mathbf{F}^{th}_{n} + \frac{c\nabla\rho_n(x)}{\rho^*} \tag{10}$$

where the prefactor c has the units of energy and scales the magnitude of the force, ρ^* is the reference density, and ρ_n is the density profile computed at step n of the iterative procedure. The calculation and application of this force has to be iterative because a single step will not be sufficient to flatten the density profile; however, the convergence to a uniform density is guaranteed by the fact that the scheme has a fixed point when $\nabla\rho = 0$. We note, in passing, that the sum of the contributions obtained from Eqs. (9) and (10) provides a force whose integral corresponds to the difference of the chemical potentials between the AT and CG domains. This method not only ensures the same density in the two subdomains, but also leads to a flat density profile also in HY region [28,29,33].

In general, the procedure to compute the thermodynamic force consists in an equilibration phase of the simulation setup where no compensation is applied (with the possible exception of the drift force compensation), followed by a production run during which an accurate density profile is computed. The latter has to be sufficiently smooth so to employ its numerical gradient as a force in the following simulation, which will provide the new density profile and so on. When the density is deemed to be uniform within a pre-established tolerance, the iterations are interrupted, and the compensation force is given by the sum of the terms computed up to that point.

As already mentioned, this scheme has the advantage of "working by default", since the new terms of the force systematically reduce the density imbalance and the amplitude of the next correction with it. However, this simple approach necessitates a possibly very large number of full simulations employed to compute the density profiles at each iteration stage with sufficient accuracy. Here we make use of an improved strategy to compute the appropriate density compensation, which is iterative as the regular one but is performed *on the fly*, and requires substantially less time.

Also in this approach the thermodynamic force is iteratively computed as the numerical gradient of the density and applied to the molecules in the HY region. The difference lies in the fact that the measurement of the density profile is performed on a very short time interval $\Delta T = \nu\delta t$, where δt is the integration time step and ν is an integer number of the order of $\sim 10^2{-}10^3$. The force is thus obtained according to Eq. (10).

The advantage of this scheme is that the small deviations of the density from a uniform profile are immediately corrected for, and the system has no time to equilibrate into a state of substantial density imbalance. However, it is obvious that the density profile computed in the small time interval ΔT would be too noisy to make any use of its numerical gradient. The solution to this problem is to convolute the position of the center of mass of a molecule with a Gaussian function with a half-width $\sigma/2$ comparable with the typical excluded volume radius of the molecules, so that the coordinates are spread on a wider range of bins rather than a single one. One thus has that the density in the bin i covering the coordinate range $[x_i, x_{i+1}]$ in a specific

simulation frame is computed as:

$$\hat{\rho}_i = \sum_\alpha \frac{1}{A} \int_{x_i}^{x_{i+1}} dy \ \exp\left[-\frac{(y-x_\alpha)^2}{2\sigma^2}\right], \tag{11}$$
$$A = \int_{-l}^{l} dy e^{-\frac{y^2}{2\sigma^2}}.$$

The parameter l, whose appropriate value depends on the specific system under examination, controls the range of the Gaussian function; a sensible choice is to set $l = 2.5\sigma$.

2.2 Reaction field

An alternative method to the Ewald summation scheme, aiming at treating electrostatics interactions, assumes a homogeneous polar fluid beyond a cutoff sphere enclosing an atom i. The charge distribution within the cavity polarizes such a fluid, and this polarization in turn influences the charge in the sphere (reaction field) [38,50]. The Coulomb potential is modified as:

$$V_{RF}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \left[1 + \frac{\epsilon_{RF}-1}{2\epsilon_{RF}+1}\left(\frac{r_{ij}}{r_c}\right)^3\right], \tag{12}$$

where r_{ij} is the inter-atomic distance, q the electric charge, ϵ_0 vacuum permittivity and r_c the cutoff distance. This expression depends on knowing beforehand macroscopic information of the system, namely, its permittivity ϵ_{RF}. Discontinuous jumps in energy occur when particles enter/leave the sphere of another particle. To tackle this problem, expression (12) has to be attenuated to zero near the cutoff radius. In particular [51]:

$$V_{RF}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \left[1 + \frac{\epsilon_{RF}-1}{2\epsilon_{RF}+1}\left(\frac{r_{ij}}{r_c}\right)^3\right] - \frac{q_i q_j}{4\pi\epsilon_0 r_c} \frac{3\epsilon_{RF}}{2\epsilon_{RF}+1}. \tag{13}$$

Finally, the force acting on atom i, derived from eq. (13), reads:

$$\mathbf{F}_{RF}(\mathbf{r}_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0}\left[\frac{1}{r_{ij}^2} - 2\frac{1}{r_c^3}\frac{\epsilon_{RF}-1}{2\epsilon_{RF}+1} r_{ij}\right]\frac{\mathbf{r}_{ij}}{r_{ij}}. \tag{14}$$

The RF method has been extensively used and both its advantages and drawbacks have been widely recognized (for a review see Ref. [41]). So far, it has been the method of choice for adaptive resolution simulations, mostly for practical reasons. However, we find that conditions such as the implicit homogeneity of the system required to describe the neighborhood of every atom in terms of a dielectric function, or the fact that it might be necessary to modify the force field to reach the desired accuracy, limit substantially the number of systems we are able to simulate. For such reasons, we turn our attention to a different method to deal with electrostatic interactions.

2.3 Damped shifted force potential (DSF)

The idea behind the DSF method was introduced in Refs. [42,52], where it was demonstrated that for a perfect ionic crystal the effective Coulomb interactions are

short ranged. Moreover, when comparing a straight cutoff method with Ewald calculations, electrostatic energies are very accurate for characteristic system-dependent cutoff distances. This tendency is due, in addition to a damped oscillatory behavior, to an almost exact charge neutrality for such particular cutoff spheres.

By combining short-range nature and charge neutrality, a pairwise summation method was introduced [42] and shown to give comparable results to standard ES. However, the use of this method is dubious for molecular dynamics simulations, in particular because of force discontinuities at the cutoff radius. To solve this problem, the DSF method was modified so to give continuous potential and forces everywhere [43], thus becoming a valuable short-range alternative to ES. In DSF, the electrostatic potential between two charges q_i and q_j separated by a distance r_{ij} is given by the following expression:

$$V_{\mathrm{DSF}}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0}\left[\frac{\mathrm{erfc}(\alpha r_{ij})}{r_{ij}} - \frac{\mathrm{erfc}(\alpha r_c)}{r_c} + \left(\frac{\mathrm{erfc}(\alpha r_c)}{r_c^2} + \frac{2\alpha}{\pi^{1/2}}\frac{\exp\left(-\alpha^2 r_c^2\right)}{r_c}\right)(r_{ij} - r_c)\right], \tag{15}$$

where $r_{ij} \leq r_c$, ϵ_0 is the vacuum permittivity, r_c is the cut-off radius, and α is a damping parameter with dimension of inverse length. $\mathrm{erfc}(r)$ is the complementary error function that takes into account the damping proposed in [42]. The gradient of potential (15) gives the force acting on atom i

$$\mathbf{F}_{\mathrm{DSF}}(\mathbf{r}_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0}\left[\frac{\mathrm{erfc}(\alpha r_{ij})}{r_{ij}^2} + \frac{2\alpha}{\pi^{1/2}}\frac{\exp(-\alpha^2 r_{ij}^2)}{r_{ij}} - \frac{\mathrm{erfc}(\alpha r_c)}{r_c^2} - \frac{2\alpha}{\pi^{1/2}}\frac{\exp(-\alpha^2 r_c^2)}{r_c}\right]\frac{\mathbf{r}_{ij}}{r_{ij}}. \tag{16}$$

We emphasise here that the short range character of electrostatic interactions has been confirmed by *ab initio* simulations of water [53]. Moreover, DSF has been successfully applied in simulations where the Ewald method can introduce spurious electrostatic effects [54]. This is the case for polarized systems when dipole-dipole interactions in the simulation box and its replicas are artificially introduced.

3 Computed quantities

Our aim is to validate the effectiveness of the DFS method to accurately reproduce the electrostatic interaction experienced by water molecules in the liquid phase, and to show that this approach is perfectly suited to be employed in the framework of an adaptive dual-resolution simulation.

To this end, we will perform a number of analysis of different structural, thermodynamical, and dynamical properties of the liquid, namely: radial distribution functions (RDFs), tetrahedral orientation order parameter, fluctuations of the number of molecules, and velocity autocorrelation functions (VACF). The results obtained in the dual-resolution setup are compared with the same quantities computed in fully atomistic benchmark simulation. In the following, a succinct description of the quantities under examination is provided.

The orientational order parameter is defined for oxygen atoms in water as [55]:

$$q = 1 - \frac{3}{8}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\left(\cos\psi_{jk} + \frac{1}{3}\right)^2, \tag{17}$$

where for a given oxygen atom i one identifies its four nearest neighbors, and computes the angles ψ_{jk} with vertex i and segments ij and ik. For a single molecule $-3 < q < 1$. By contrast, for a collection of molecules $0 \leq \langle q \rangle < 1$, with 0 corresponding to the ideal gas case and 1 to a perfect tetrahedral network.

Fluctuations of the number of molecules are calculated using the expression:

$$\left.\frac{\Delta^2(N)}{\langle N \rangle}\right|_x = \left.\frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle}\right|_x \tag{18}$$

where the subscript x indicates that the simulation box has been divided in slabs of width 10 Å along the X axis. The average $\langle N \rangle$ and standard deviation $\Delta^2(N)$ in the number of molecules have been calculated for such slabs.

Finally, the velocity autocorrelation function (VACF) is defined as [50]:

$$C_{vv}(t) = \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle \,, \tag{19}$$

where $\mathbf{v}_i(t)$ is the velocity of molecule i at time t. To integrate numerically Newton's equations of motion, molecular dynamics simulations rely on discrete time algorithms. Therefore, Eq. (19) should be estimated to take into account such discretization. Here we use the discrete estimator described in Ref. [56] where the VACF for the t_m-th time step takes the form:

$$C_{vv}(t_m) = \frac{1}{N_{AA}} \sum_{i=1}^{N_{AA}} \frac{1}{M-m} \sum_{n=0}^{M-m-1} \mathbf{v}_i(t_{n+m}) \cdot \mathbf{v}_i(t_n) \,, \tag{20}$$

with M the total number of time steps and $t_m = m\delta t$, where δt is the integration time step. This expression is constructed in such a way that it includes all possible contributions $\mathbf{v}(t + n\delta t) \cdot \mathbf{v}(t)$ that result from shifting the time origin by m steps. The normalization factor $1/(M-m)$ ensures that the estimator is unbiased. We implemented Eq. (20) following the protocol reported in Ref. [57]. Finally, N_{AA} is the number of molecules that always remain within a predefined region of the simulation box. In the case of fully-atomistic simulations, $N_{AA} = N$, the total number of molecules. The error in the calculation of the VACF is given by $2t_{corr}/N_{AA}t_{tot}$ [50], with t_{corr} the correlation time and t_{tot} the total time of the simulation.

4 Simulation details

In all case studies unless otherwise stated, there are 10240 water molecules in the simulation box. The time step for the `H-AdResS` as well as the fully atomistic simulations is set to $\delta t = 1\,\text{fs}$. The initial configuration for every simulation setup is extracted from the simulation results of a similar fully atomistic system which is equilibrated for 50 ps in the isothermal-isobaric ensemble at a temperature $T_0 = 300\,\text{K}$ and pressure $P_0 = 1\,\text{bar}$. The atomistic interactions between the atoms of water molecules are based on the SPC/E model [58–60].

Three different methods have been used to simulate the atomistic Coulomb interactions: particle-particle-particle-mesh (PPPM) Ewald summation [48], reaction field (RF) [37,38], and damped shifted potential (DSF) [42,43]. For the latter, everywhere in the present work we used the following parameters, for which the DSF potential best reproduces the RDFs of the reference ES simulations: damping parameter $\alpha = 0.2\,\text{Å}^{-1}$; cut-off radius $r_c = 12\,\text{Å}$.

The Weeks-Chandler-Andersen (WCA) [61] potential is used for the interaction of water molecules in the coarse-grained domain:

$$V^{CG}(r) = \begin{cases} 4\epsilon \left[(\sigma/r)^{12} + (\sigma/r)^{6} \right] & r \leq 2^{1/6}\sigma \\ 0 & r > 2^{1/6}\sigma \end{cases} \tag{21}$$

where $\epsilon = 1.0$ kcal/mol and $\sigma = 2.2$ Å, which is roughly the excluded volume diameter of the water molecules in the fully atomistic simulations.

In all `H-AdResS` cases, the initial equilibrated configuration is simulated for 100 ps in the Canonical (NVT) ensemble; A uniform temperature profile at 300 K is maintained throughout the system (data not shown) via a Nosé-Hoover thermostat with a damping parameter of 0.1 ps. In this step, the equations of motion are integrated according to the Hamiltonian in Eq. (1) in absence of the compensation terms. The presence of the drift force in the HY region leads to a pressure imbalance between two resolutions (see solid blue curve in Fig. 5). After the initial 100 ps, the *on-the-fly* calculation of the drift force correction is applied. The resolution range is subdivided in 20 bins of size $\Delta\lambda = 0.05$. The drift force compensation is updated every 1 ps for 150 ps. At the end of this step, the converged compensation force ultimately cancels out the drift force and leads to a uniform pressure across the simulation domains (see red dashed line in Fig. 5), while a density gradient persists across the HY region. To enforce a uniform density the *on-the-fly* density balancing method is applied for the next 300 ps. In this step, the length of the simulation box is uniformly discretized into slabs of size $\Delta x = 1.5$ Å and the thermodynamic force is updated every 0.5 ps. We employed values of $c = 2.0$, $\sigma = 6$ Å and $l = 12$ Å for smoothing and scaling the thermodynamic force.

All simulations are performed with the `LAMMPS` simulation package, with the exception of the RF runs, that have been performed on the `GROMACS` [51] platform.

5 Results and discussion

In this section we report the results of the validation of the DSF method for the electrostatic interaction in the context of a `H-AdResS` simulation. The first part is devoted to the comparison of the two short-range modifications of Coulomb potential (RF and DSF) with the Ewald summation PPPM scheme, that we take here as our golden standard. Subsequently, we focus on the DSF method and compare the properties of water in the AT region of the `H-AdResS` setup with those measured in an equivalent domain of a reference fully atomistic simulation.

5.1 All atom simulations

The three methods here under exam to reproduce electrostatic interactions in MD simulations have been extensively investigated in the past. However, we consider useful to include this simulations here to provide a self-contained validation of their performance within the framework of the `H-AdResS` scheme.

From the point of view of structural properties, PPPM, DSF and RF give identical results. In particular, RDFs for the three cases can be seen in Fig. 2 where the three data sets overlap perfectly.

This is also the case for the orientational order parameter q. Figure 3 shows the normalised distribution of q for the three cases considered where, as expected, a bimodal character is observed [55]. In such a distribution, low values of q, related to angular distorsions, indicate local disorder of water molecules. By contrast, since q is

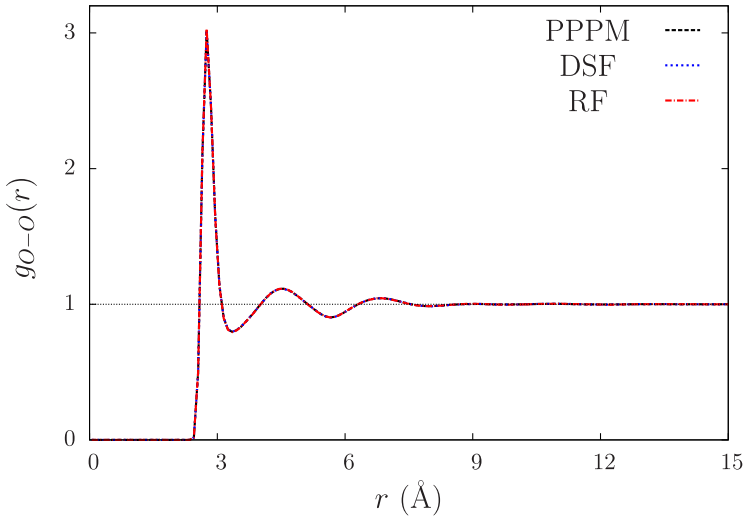

Fig. 2. Oxygen-Oxygen radial distribution function (RDF), $g_{O-O}(r)$, for SPC/E water using three different approaches to compute electrostatic interactions. Namely, particle-particle particle-mesh (PPPM), damped shifted force potential (DSF) and reaction field (RF).

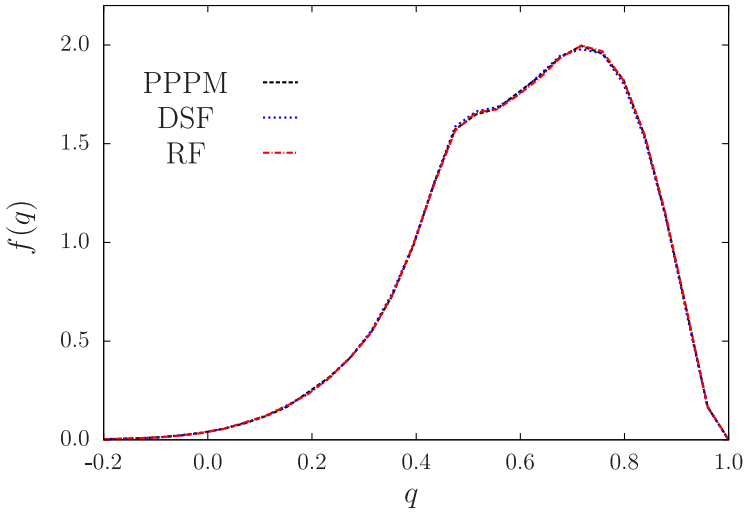

Fig. 3. Normalised distribution of the orientational order parameter. The fraction of molecules with a given value q is given by $f(q)dq$.

associated to the angular ordering of first neighbours and ignore their radial ordering, high values of q do not necessarily demonstrate a tetrahedral ordering [62].

The short-time dynamics of water molecules has been investigated by means of the oxygen VACF. In all three cases under examination we have run a 2 ps long simulation in absence of the thermostat, with a time step of 1 fs and recording velocities every 10 fs. The error in the VACF is $\sim 1\%$, and it is estimated using $2t_{corr}/N_{AA}t_{tot}$ with $t_{corr} = 1$ ps [50]. Consistently with the structural results, also the dynamical (equilibrium) behavior of the system is not affected by the different method employed to treat Coulomb interaction. A subtle deviation of the RF from the other two setups is to be attributed to numerical discrepancies due to the different software (`GROMACS` [51]) employed in the former case.

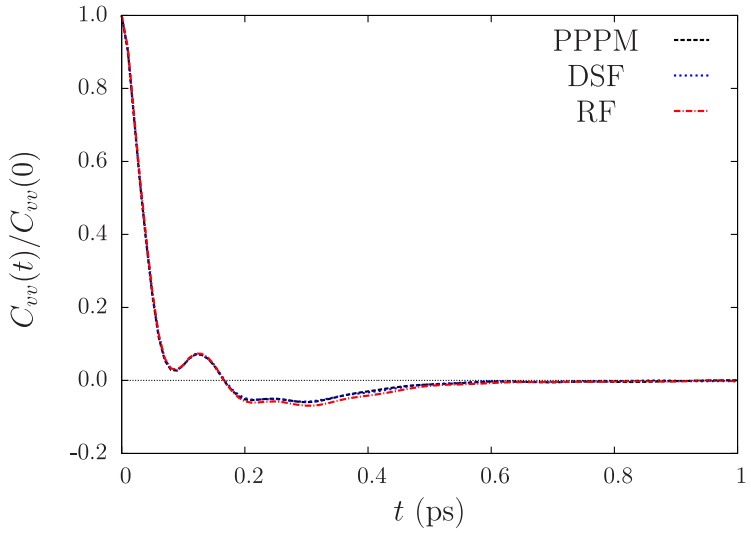

Fig. 4. Normalized velocity autocorrelation function, $C_{vv}(t)/C_{vv}(0)$, for SPC/E water using three different approaches to compute electrostatic interactions, as in Fig. 2.

The reported analysis shows that the DSF scheme is capable of reproducing quantitatively accurate structural and dynamical properties of liquid water, in addition to ionic liquids and other complex systems [43,63–67]. The method is thus a strong candidate to replace the RF as the "short-range alternative" to Ewald summation in dual-resolution simulation schemes.

5.2 H-AdResS simulations

The most basic requirement of an adaptive dual-resolution simulation is that the compensation applied to the molecules in the HY region is sufficient to enforce a uniform density profile of the fluid across the simulation domain. That this is the case in the water model under examination is confirmed by the data reported in Fig. 5, which provide a clear picture of thermodynamic properties of the `H-AdResS` setup in terms of pressure and density profiles.

Without any compensation (solid blue line Fig. 5) the system equilibrates in such a way that both pressure and density (upper and lower panel, respectively) are different in the AT and CG regions, and therefore different from the reference values. The situation changes if the drift force is counterbalanced, or in other words, a pressure correction is included which removes on average the drift force. The profiles in this case show that the pressure is now the same in both subregions; the density has improved thanks to the removal of the extra pressure exerted on the molecules in the HY region, however it is still higher in the AT region (red dashed line Fig. 5).

Finally, after the application of the thermodynamic force to compensate the density imbalance (dotted green line Fig. 5), we observe that in this setup the pressure is different in the two bulk subdomains whereas the density is uniform, with a deviation of one percent from the reference value. In the following, we analyze our results for the case where both compensations are applied.

From the structural point of view, the fully atomistic and `H-AdResS` simulations provide perfectly compatible results. This is evident from Fig. 6, where we report both sets of multicomponent RDFs which exhibit an excellent overlap. The normalized

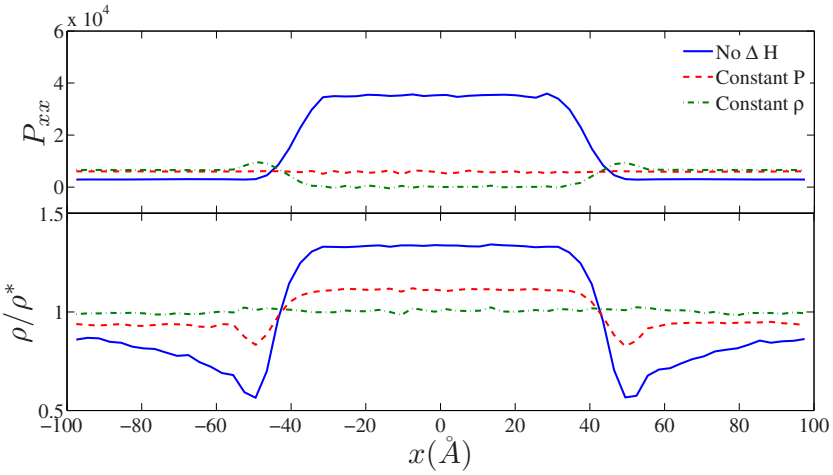

Fig. 5. Pressure (top) and density (bottom) profiles for different `H-AdResS` setup. The blue (solid) curves represent `H-AdResS` setup with no compensation in the Hamiltonian. The red dashed curves illustrate the setup with a constant-pressure route, and the results of the setup with constant-density route are shown in green dash-dotted curves.

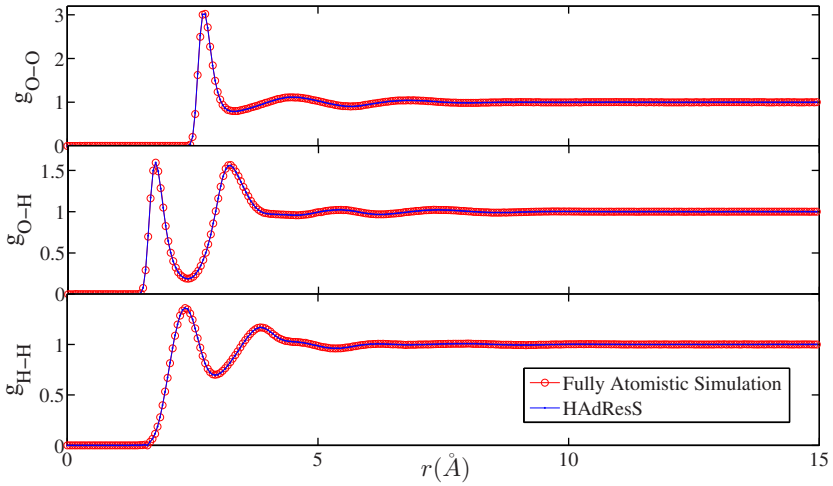

Fig. 6. RDFs of water molecules at pressure $P_0 = 1\,\text{bar}$ and temperature $T_0 = 300\,\text{K}$ in two different simulation setups: fully atomistic simulation (red line with open circles) and `H-AdResS` (blue line with dots). From top to bottom, the plots show oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen RDFs. The DSF damping parameter is set to $\alpha = 0.2\,\text{Å}^{-1}$, and the cut-off radius is $R_c = 12\,\text{Å}$.

distributions of the orientational order parameter for the fully atomistic and dual-resolution cases are shown in Fig. 7, and they clearly overlap with great accuracy. The bimodal profile, observed in both cases, indicates that the tendency of water molecules to form ordered structures is well preserved in `H-AdResS` simulations.

A relevant property that has to be correctly reproduced in the AT region in order to guarantee that the thermodynamics of this subdomain is representative of the reference simulation is the profile of fluctuations of the number of particles [28,68]. Profiles of fluctuations for fully atomistic and `H-AdResS` simulations, reported in Fig. 8, coincide in the AT subdomain. Beyond the hybrid region, as expected, the

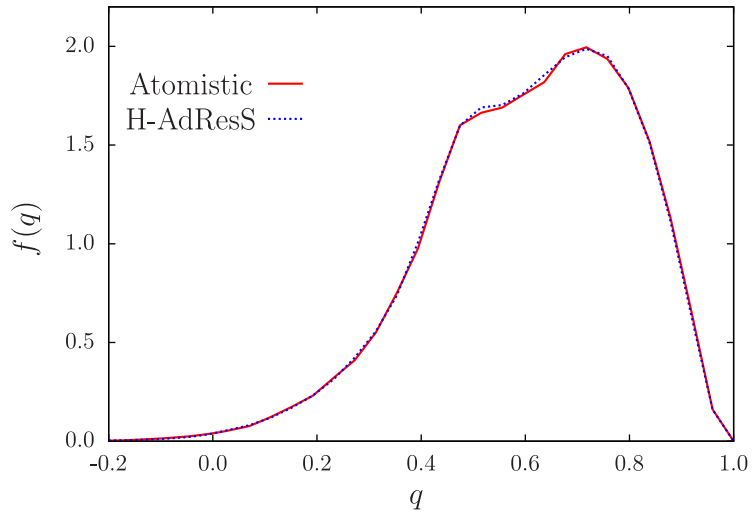

Fig. 7. Normalized distribution of the orientational order parameter for fully-atomistic and `H-AdResS` simulations. Only water molecules in the interval $-25\,\text{Å} < x < 25\,\text{Å}$ were considered for the calculation.

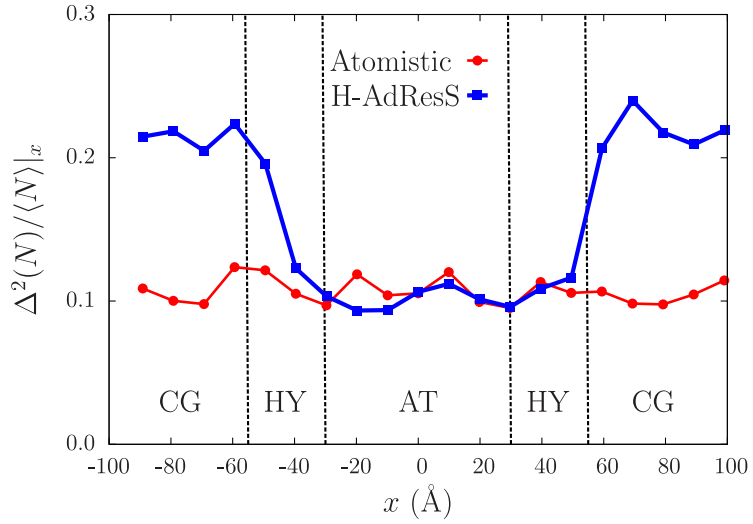

Fig. 8. Profile of fluctuations of the number of molecules calculated along the x-axis. Red circles correspond to a benchmark fully atomistic simulation of water molecules. Blue squares correspond to an `H-AdResS` simulation of a slab of water molecules enclosed by a WCA system. Fluctuations are calculated using sub-volumes of size $10 \times 40 \times 40\,\text{Å}$. Vertical lines are guides to the eye and indicate the location of the AT, HY, and CG regions of the `H-AdResS` setup.

profile of fluctuations increases, due to the different isothermal compressibility of the CG model.

Finally, concerning short-term dynamical properties, we confirm that fully atomistic and `H-AdResS` simulations display consistent VACFs. The measurement has been performed only in the AT region in both cases, hence the error in the VACF ($\sim 3\%$) is higher than in the measurement performed for the fully atomistic cases with different electrostatics (Fig. 4). Nonetheless, differences observed in both fully atomistic

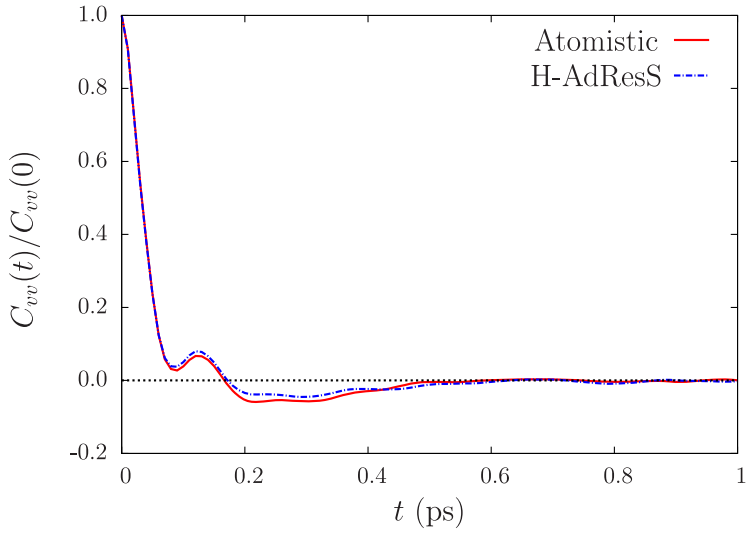

Fig. 9. Normalized velocity autocorrelation function for fully-atomistic (full red line) and H-AdResS (dashed blue line) simulations of a 2 ps trajectory with a time step of 1 fs. Water molecules in the interval $-25\,\text{Å} < x < 25\,\text{Å}$ were considered for the calculation.

and `H-AdResS` cases are larger than error bars ($\sim 10\%$) in the 0.2−0.4 ps interval. To explain this aspect, let us recall that in `H-AdResS` simulations as presented here, the density in the atomistic region is approximately 1% above the reference density. Therefore, differences in VACF appear because we compare systems with slightly different densities, as indicated by further fully atomistic simulations performed at higher density (results not shown). Interestingly, this strong density dependence of the VACF is the matter of recent discussion [69].

6 Conclusions

The appropriate treatment of electrostatic interaction in computer simulations of soft and biological matter is still an open problem. One of the most challenging issues is the possibility to find a balance between the accurate description of the potential and a computationally economic implementation of the corresponding model. If, on the one hand, the original decomposition of the interaction in short- and long-range terms devised by Ewald has been substantially optimized, on the other hand some of the undesired artifacts due to the unphysical periodicity implicit in this treatment are still cause of concern in specific systems. In the context of adaptive, dual-resolution simulations Ewald summation-based schemes are in any case practically unviable, as the long-range term would have to be computed on models featuring substantially different physical properties. Alternative modifications of Coulomb potential, such as the reaction field approach, circumvent these problems and provide a computationally effective short-range interaction; also in this case, however, there are some limitations originating in the underlying assumption of a uniform medium, which do not always apply.

The DSF potential, on the other hand, has been shown to reproduce the physical (structural and dynamical) properties of many charged systems with high accuracy, albeit being short-range and without the necessity of a pre-parametrization based on emergent properties of the system, e.g. the dielectric constant. This method is thus

ideally suited to be employed in the context of dual-resolution simulations, and its validity has been here demonstrated by means simulations of liquid water. Specifically, the region of the dual-resolution setup where the fluid was modeled with atomistic resolution showed quantitatively consistent properties compared to a reference, fully atomistic simulation.

The possibility to accurately reproduce the effects of Coulomb potential in the context of adaptive resolution simulations without the need to parametrize the atomistic force field, as it would be the case when employing the reaction field, thus opens the way to the efficient modeling and simulation of complex systems in which electrostatic interaction is known to play a primary role, for example ions in solutions, ionic liquids, and nucleic acids. Additionally, the flexible and efficient implementation of the `H-AdResS` method in the `LAMMPS` simulation package, equipped with the DSF method for Coulomb potential, provides a broad community with an effective instrument to investigate soft and biological matter.

Appendix: LAMMPS implementation

We report here the basic technical details of the `H-AdResS` implementation in the `LAMMPS` simulation package. The software as well as a more detailed documentation can be downloaded at the web page: `http://www2.mpip-mainz.mpg.de/~potestio/software.php` or cloned from the `LAMMPS` Git repository `https://github.com/lammps/lammps`.

6.1 `H-AdResS` Atom style

We introduced an atom style called `full/hars` in which an atom i, in addition to `LAMMPS`'s indigenous atom properties (e.g. coordinates $\mathbf{x}_i$, velocities $\mathbf{v}_i$, charge $\mathbf{q}_i$), is provided the following `H-AdResS`-specific attributes:

- λ_i: the resolution of the atom in the system as determined by the value of the switching function computed on the center of mass coordinate of the molecule to which atom i belongs.
- $\nabla\lambda_i$: the gradient of the switching function.
- $\mathbf{x}_\alpha^{CG}$: the center of mass coordinate of molecule α to which atom i belongs.
- $\texttt{Rep}_i$: the *representation* flag indicating which atom in the molecule carries the information pertaining the whole molecule.
- $\texttt{MolType}_i$: the molecule type index specifying the CG model parameters of molecule when in the low resolution region.

All these properties are assigned and set by the two files `atom_vec_full_hars.cpp/h`.

6.2 Fixing particle resolutions

As the atom passes through different resolutions, the resolution function λ_i and its gradient $\nabla\lambda_i$ have to be updated at each time step. This step is carried out within a fix file called `fix_lambdah_calc.cpp/h`: here, the center of mass coordinates of each molecule, $\mathbf{x}_\alpha^{CG}$, is calculated and then spread to all atoms of that molecule. Depending on the (user-specified) shape of the hybrid region, also the switching function and its gradients are computed based on the molecule position and transmitted to the corresponding atoms.

6.3 Calculating coarse-grained pairwise interaction

The generalized pairwise coarse-grained potentials and forces are computed based on Eqs. (3) and (4). Depending on the type of potential, a specific interaction file (with corresponding header) is introduced. For this study, the interaction between the water molecules in the CG region is given by a WCA potential which, in turn, is obtained by assigning a specific set of parameters to a Lennard-Jones potential; accordingly, the two files `pair_lj_cut_hars_cg.cpp/h` overseeing the computation of Lennard-Jones interaction in the CG and HY regions are employed. The computation of the forces among neighboring molecules is restricted to the representative atoms provided with the details about the molecule properties. Hence, for molecule α the coarse-grained force due to the interaction with all other neighboring β molecules is computed as:

$$\mathbf{F}_\alpha = \sum_{\beta,\beta\neq\alpha} \left\{ \left(1 - \frac{\lambda_\alpha + \lambda_\beta}{2} \right) \mathbf{F}^{CG}_{\alpha|\beta} \right\} + V^{CG}_\alpha \nabla_\alpha \lambda_\alpha. \tag{22}$$

The force $\mathbf{F}_{\alpha i}$ acting on atom i^{th} is obtained by scaling the molecular force $\mathbf{F}_\alpha$ by each atom's mass m_i divided by the whole molecule's mass M_α:

$$\mathbf{F}_{\alpha i} = \frac{m_i}{M_\alpha} \mathbf{F}_\alpha. \tag{23}$$

6.4 Calculating atomistic pairwise interaction

A procedure similar to the one discussed for the CG potential is carried out for the atomistic part of the interaction. Depending on the specific atomistic force field, two files need to be written and added to `LAMMPS`'s source directory. For our study we employed Lennard-Jones and DSF Coulomb interactions, and the files `pair_lj_cut_coul_dsf_hars_at.cpp/h` have been accordingly created. The pairwise atomistic interactions between neighboring atoms are calculated through:

$$\mathbf{F}_{\alpha i} = \sum_{j\in\beta, i\in\alpha, i\neq j} \frac{\lambda_\alpha + \lambda_\beta}{2} \mathbf{F}^{AT}_{\alpha i|\beta j}. \tag{24}$$

Since drift forces are acting on the molecules, an additional force contribution is added to each molecule:

$$\mathbf{F}_\alpha = -\left(\sum_{i\in\alpha} V^{AT}_{\alpha i} \right) \nabla_\alpha \lambda_\alpha. \tag{25}$$

Afterwards, the computed molecular drift force is spread to the atoms of molecule α based on Eq. (23). For the intra-molecular interactions, such as bond and angle potentials, there is no need to modify the corresponding parts of the current `LAMMPS` implementation.

6.5 Speedup

The reduced number of interactions in the CG region and their shorter range enable a reduction of the computational cost of the simulation. In order to quantitatively demonstrate this gain, we have performed fully atomistic as well as dual-resolution simulations of water systems of increasing size, and compared their run time. In Fig. 10

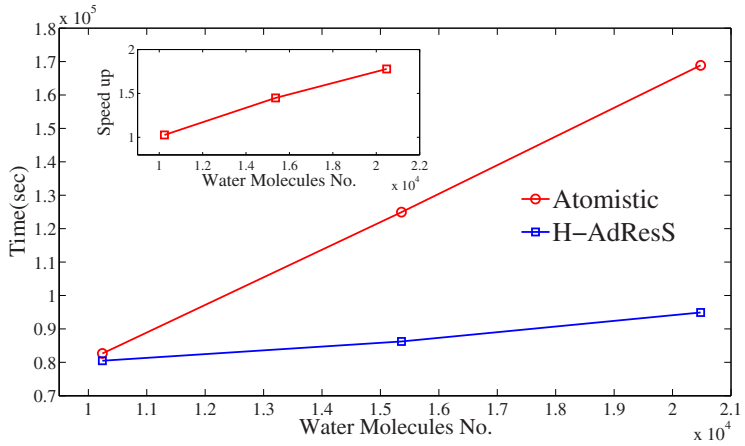

Fig. 10. Computation time in seconds for fully atomistic simulations (red circles) and `H-AdResS` (blue squares) is shown for three different simulation box sizes. For the case of `H-AdResS`, the length of the atomistic and hybrid region are kept constant as the length of the simulation box increases. The speedup which is calculated by the computation time ratio of fully atomistic simulation to `H-AdResS` is shown in onset.

we show the time necessary to these different setups to perform the same number of integration steps, namely 10^5, and the corresponding speedup, calculated as the ratio of the time of a fully atomistic run over the corresponding dual-resolution one. The simulated system is the same discussed in the Methods section; the AT interactions are given by DSF and Lennard-Jones potentials with a cutoff radius of 12 Å, and the CG model is a purely repulsive WCA. The widths of the AT and HY regions are kept constant, while the size of the CG domains is systematically increased; the probed extensions of the CG domain are thus approximately 90, 180, and 270 Å. By comparing the simulation time of fully atomistic and dual-resolution setups it is possible to appreciate that the latter has a very weak linear growth as a function of the CG domain size, indicating that the computational cost of the simulation is almost completely determined by the computation of AT interactions.

MH, DD and RP acknowledge financial support under the project SFB-TRR146 of the Deutsche Forschungsgemeinschaft. RCH acknowledges financial support from the Alexander von Humboldt Foundation. The authors are thankful to Kostas Daoulas for an attentive reading of the manuscript and useful comments. The authors are also deeply grateful to Kurt Kremer for years of fruitful collaborations and scientific discussions, in particular on the topic of adaptive resolution simulations.

References

1. G. Grest, K. Kremer, Phys. Rev. A. **33**, 3628 (1986)
2. K. Kremer, G. Grest, I. Carmesin, Phys. Rev. Lett. **61**, 566 (1988)
3. L. Yelash, M. Müller, W. Paul, K. Binder, J. Chem. Theory Comput. **2**, 588 (2006)
4. T. Spyriouni, C. Tzoumanekas, D. Theodorou, F. Müller-Plathe, G. Milano, Macromolecules **40**, 3876 (2007)
5. J. McCammon, M. Karplus, Nature **268**, 765 (1977)
6. M. Karplus, J. McCammon, Nature **277**, 578 (1979)
7. P. Raiteri, A. Laio, F.L. Gervasio, C. Micheletti, M. Parrinello, J. Phys. Chem. B. **110**, 3533 (2006)

8. H. Lou, R.I. Cukier, J. Phys. Chem. B. **110**, 12796 (2006)
9. K. Arora, C.L. Brooks, Proc. Natl. Acad. Sci. USA **104**, 18496 (2007)
10. F. Pontiggia, A. Zen, C. Micheletti, Biophys. J. **95**, 5901 (2008)
11. M.M. Tirion, D. ben Avraham, J. Mol. Biol. **230**, 186 (1993)
12. M.M. Tirion, Phys. Rev. Lett. **77**, 1905 (1996)
13. I. Bahar, A.R. Atilgan, B. Erman, Folding and Design **2**, 173 (1997)
14. C. Micheletti, P. Carloni, A. Maritan, Proteins **55**, 635 (2004)
15. R. Potestio, F. Pontiggia, C. Micheletti, Biophys. J. **96**, 4993 (2009)
16. C. Globisch, V. Krishnamani, M. Deserno, C. Peter, PLoS. ONE **8**, e60582, 04 (2013)
17. K. Kremer, Comput. Simul. Soft Matter Sci. **53** (2000)
18. K. Kremer, F. Müller-Plathe, MRS Bull. **26**, 205 (2001)
19. N.A. van der Vegt, C. Peter, K. Kremer, *Structure-Based Coarse- and Fine-Graining in Soft Matter Simulations* (CRC Press - Taylor and Francis Group, 2009)
20. C. Hijón, E. Vanden-Eijnden, R. Delgado-Buscalioni, P. Español, Farad. Discuss. **144**, 301 (2010); discussion 323–45, 467–81 (2010)
21. W. Noid, *Systematic Methods for Structurally Consistent Coarse-Grained Models*, **924** of *Methods in Molecular Biology* (Humana Press, 2013)
22. W.G. Noid, J. Chem. Phys. **139**, 090901 (2013)
23. M. Praprotnik, L. Delle Site, K. Kremer, J. Chem. Phys. **123**, 224106 (2005)
24. M. Praprotnik, L. Delle Site, K. Kremer, Phys. Rev. E **73**, 066701 (2006)
25. M. Praprotnik, L. Delle Site, K. Kremer, J. Chem. Phys. **126**, 134902 (2007)
26. M. Praprotnik, L. Delle Site, K. Kremer, Ann. Rev. Phys. Chem. **59**, 545 (2008)
27. S. Fritsch, C. Junghans, K. Kremer, J. Chem. Theo. Comput. **8**, 398 (2012)
28. S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, K. Kremer, Phys. Rev. Lett. **108**, 170602 (2012)
29. R. Potestio, S. Fritsch, P. Español, R. Delgado-Buscalioni, K. Kremer, R. Everaers, D. Donadio, Phys. Rev. Lett. **110**, 108301 (2013)
30. R. Potestio, P. Español, R. Delgado-Buscalioni, R. Everaers, K. Kremer, D. Donadio, Phys. Rev. Lett. **111**, 060601 (2013)
31. A. Agarwal, H. Wang, C. Schütte, L.D. Site, J. Chem. Phys. **141**, 034102 (2014)
32. K. Kreis, D. Donadio, K. Kremer, R. Potestio, Europhys. Lett. **108**, 30007 (2014)
33. P. Español, R. Delgado-Buscalioni, R. Everaers, R. Potestio, D. Donadio, K. Kremer, J. Chem. Phys. **142**, 064115 (2015)
34. B.J. Lambeth, C. Junghans, K. Kremer, C. Clementi, L. Delle Site, J. Chem. Phys. **133**, 221101 (2010)
35. A.C. Fogarty, R. Potestio, K. Kremer, J. Chem. Phys. **142**, 195101 (2015)
36. P.P. Ewald, Ann. Phys. **64**, 253 (1921)
37. L. Onsager, J. Amer. Chem. Soc. **58**, 1486 (1936)
38. J.A. Barker, R.O. Watts, Mol. Phys. **26**, 789 (1973)
39. D. van der Spoel, P.J. van Maaren, H.J.C. Berendsen, J. Chem. Phys. **108**, 10220 (1998)
40. S. Bevc, C. Junghans, K. Kremer, M. Praprotnik, New J. Phys. **15**, 105007 (2013)
41. I. Fukuda, H. Nakamura, Biophys. Rev. **4**, 161, (2012)
42. D. Wolf, P. Keblinski, S.R. Phillpot, J. Eggebrecht, J. Chem. Phys. **110**, 8254 (1999)
43. C.J. Fennell, J.D. Gezelter, J. Chem. Phys. **124**, 23 (2006)
44. M.A. Kastenholz, P.H. Hünenberger, J. Phys. Chem. B **108**, 774 (2004)
45. D.A.C. Beck, R.S. Armen, V. Daggett, Biochemistry **44**, 609 (2005)
46. R.D. Lins, U. Röthlisberger, J. Chem. Theo. Comput. **2**, 246 (2006)
47. S. Plimpton, J. Comput. Phys. **117**, 1 (1995)
48. R.W. Hockney, J.W. Eastwood, *Computer simulation using particles* (CRC Press, 1988)
49. J. Kirkwood, J. Chem. Phys. **3**, 300 (1935)
50. M.P. Allen, D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987)
51. D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, J. Comp. Chem. **26**, 1701 (2005)
52. D. Wolf, Phys. Rev. Lett. **68**, 3315 (1992)

53. S. Fritsch, R. Potestio, D. Donadio, K. Kremer, J. Chem. Theory Comput. **10**, 816 (2014) PMID: 26580055
54. F. Bresme, A. Lervik, D. Bedeaux, S. Kjelstrup, Phys. Rev. Lett. **101**, 020602 (2008)
55. J.R. Errington, P.G. Debenedetti, Nature **409**, 318 (2001)
56. J. Kohanoff, Comput. Mat. Sci. **2**, 221 (1994)
57. T. Youngs, "*dlputils: Calculate properties from molecular dynamics trajectories*," (2016) `https://www.projectaten.com/dlputils`
58. H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, J. Phys. Chem. **91**, 6269 (1987)
59. L.X. Dang, B.M. Pettitt, J. Phys. Chem. **91**, 3349 (1987)
60. Y. Wu, H.L. Tepper, G.A. Voth, J. Chem. Phys. **124**, 2 (2006)
61. J.D. Weeks, D. Chandler, H.C. Andersen, J. Chem. Phys. **54**, 5237 (1971)
62. E. Duboué-Dijon, D. Laage, J. Phys. Chem. B, **119**, 8406 (2015) PMID: 26054933
63. C. Avendaño, A. Gil-Villegas, Mol. Phys. **104**, 1475 (2006)
64. T.G. Desai, J. Chem. Phys. **127** (2007)
65. Y. Nagata, S. Mukamel, J. Amer. Chem. Soc. **132**, 6434 (2010)
66. E.E. Gdoutos, R. Agrawal, H.D. Espinosa, Inter. J. Numer. Meth. Eng. **84**, 1541 (2010)
67. W. Shi, E.J. Maginn, J. Phys. Chem. B **112**, 2045 (2008)
68. K. Kreis, A.C. Fogarty, K. Kremer, R. Potestio, Eur. Phys. J. Special Topics **224**, 2289 (2015)
69. S. Bellissima, M. Neumann, E. Guarini, U. Bafile, F. Barocchi, Phys. Rev. E **92**, 042166, (2015)