Regular Article

Concurrent coupling of realistic and ideal models of liquids and solids in Hamiltonian adaptive resolution simulations[⋆]

Maziar Heidari[1], Robinson Cortes-Huerto[1], Kurt Kremer[1], and Raffaello Potestio[1,2,3,a]

[1] Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany
[2] Physics Department, University of Trento, via Sommarive, 14 I-38123 Trento, Italy
[3] INFN-TIFPA, Trento Institute for Fundamental Physics and Applications, I-38123 Trento, Italy

Abstract. To understand the properties of a complex system it is often illuminating to perform a comparison with a simpler, even idealised one. A prototypical application of this approach is the calculation of free energies and chemical potentials in liquids, which can be decomposed in the sum of ideal and excess contributions. In the same spirit, in computer simulations it is possible to extract useful information on a given system making use of setups where two models, an accurate one and a simpler one, are concurrently employed and directly coupled. Here, we tackle the issue of coupling atomistic or, more in general, interacting models of a system with the corresponding idealised representations: for a liquid, this is the ideal gas, *i.e.* a collection of non-interacting particles; for a solid, we employ the ideal Einstein crystal, a construct in which particles are decoupled from one another and restrained by a harmonic, exactly integrable potential. We describe in detail the practical and technical aspects of these simulations, and suggest that the concurrent usage and coupling of realistic and ideal models represents a promising strategy to investigate liquids and solids *in silico*.

1 Introduction

Since the dawn of computer-aided research, when electronic calculators were employed to numerically solve analytical expressions, a steadily growing fraction of scientific investigation relies on the help of machines [1–4]. This is particularly true for the field of soft matter [5–8], where computer simulations have enabled researchers to "infuse life" in models of increasing complexity and investigate the behaviour of systems ranging from hard spheres liquids [9] to melts of polymers [10], novel materials [11], or biomolecules such as proteins [12–16] and DNA [17]. The power of the computational approach lies especially in the flexibility and arbitrariness of the model design, where level of resolution, interaction potentials, particle-based or continuum description, thermodynamical conditions and many other features are left to the modeller, who has free hand in creating a world where her/his rules apply.

Computer simulations have thus broken the chains that kept us bound to exact results and simple approx-

imations of even simpler models, allowing us to study strongly interacting systems with several different types of potentials and large sizes [18]. The determination of free energy differences between states of a given system, for example, does not require the exact calculation of its partition function or some multi-body expansion of the latter; rather it can be (often) performed by means of standard techniques such as thermodynamic integration [19], umbrella sampling [5,6], or any other enhanced sampling algorithm [20].

And yet, the simplicity, exact solvability, or physical intuitiveness of the simplest models exert not only an intellectual attraction on the scientist fascinated by the universality of the concepts and properties that these models entail; on the contrary, they continue to represent a pivotal element in the construction of our picture of reality, in that they serve as bedrock, reference, and gauge for many other, more complex systems. One example for all: the absolute values of fundamental thermodynamical properties, such as free energies and chemical potentials, can almost always be separated in two terms: an ideal part, originating from the exactly solvable ideal contribution, and an excess part, which contains the effect of the interactions present in the "real" system. Remarkable is that these two terms are exactly additive (again, with exceptions) in a non-perturbative fashion [21,6]. In gen-

eral, then, the calculation of important system properties goes through the comparison between the most accurate model one can simulate on a computer and a simpler, more tractable representation much of which is already known. The additivity of the property under investigation does the rest.

In recent years, a peculiar class of computer simulations has emerged as a rich and versatile tool to perform model comparison, that is, adaptive resolution simulations [22–25]. These are setups in which the same systems, typically a fluid, is represented within the same simulation by means of two different models at different resolution. The simulation domain is subdivided in two parts: a high-resolution region, where the more accurate and computationally expensive model is employed, and a low-resolution region, where the simpler description featuring a lower computational cost is used. Depending on its position in space, a molecule is described by one model or the other; however, an open boundary, geometrically separating these two domains, allows the molecules to diffuse freely and change model, *i.e.* resolution, on the fly. Within a finite-sized layer located at the interface between the two main subregions, dubbed hybrid or transition region, the molecule smoothly varies its resolution, adapting its representation to one model or the other in a continuous manner.

The motivations behind this setup, in which the position in space dictates a molecule's model, are several. The simplest and most obvious is the computational gain: in the largest fraction of the simulation volume the computationally "cheaper" model is employed, thus reducing the amount of resources necessary to calculate forces and potentials; on the contrary, in a small volume —where the interesting things happen— the system is described with high accuracy, and the smooth coupling with the lower-resolution environment preserves its thermodynamical properties (*e.g.*, density, temperature, particle number fluctuations . . .).

There is a second strength of adaptive resolution simulations with many farther-reaching consequences and applications: such setups, in fact, enable one to provide, by means of the low-resolution model, a sufficiently accurate description of the long-wavelength properties of the system at the length scale of the whole simulation box, while the short-wavelength features are only locally modelled. Besides the computational gain this approach entails, one gains the opportunity of modulating the global and the local independently and, through this, to investigate their interplay [26–29]. This strategy offers clear advantages with respect to a simulation *limited to* the sole high-resolution region, as it also includes the effect of the environment at a reduced computational cost and suppresses finite-size effects.

Another important *raison d'être* of these methods is the fact that the space-dependent coupling of two different models establishes a thermodynamical relationship between them from which nontrivial information can be obtained. In fact, in adaptive resolution simulations forces emerge, whose origin depends on the specific technique employed, that steer the system towards an equilibrium

state where, *e.g.*, the local densities in the two subdomains differ. To attain the same density in all parts of the simulation one has to impose a single-molecule potential that can be traced back to the difference in Gibbs free energy between the two models concurrently employed to represent the same system [30,31]. Hence, in the process of parametrising the setup so as to have a uniform density profile, one quantifies the liquid's chemical potential difference between the simple and the accurate representation [32]. This procedure is similar in spirit to a thermodynamic integration [19], however it does not require the time-dependent switch of the Hamiltonian between two models, rather a single run concurrently contains the information about the two end states and all those in between.

It is at this point that the simplest, most fundamental models enter the scene. With a tool at hand that "compares" the chemical potentials of two systems (more precisely: two different representations of the same system) it is natural to think of employing, as a low-resolution model, one that is as inexpensive, simple, understood, and exactly solvable as possible. In the realm of fluids, such a model is the ideal gas, that is, a collection of non-interacting particles (representing atoms as well as molecules) fully described by temperature and density. All relevant thermodynamical properties of the ideal gas are known and can be calculated exactly, and represent the main contribution to those observables, such as free energies, that can be decomposed into the sum of the ideal part, determined indeed by the ideal gas, and the excess part due to the configurational partition function and the interaction potentials within it.

As adaptive resolution simulation methods enable the direct, smooth coupling of a given model to its corresponding ideal representation, the possibility opens to a new class of approaches to extract excess quantities. However, while for dense liquids such as water the coupling with an ideal gas has already been performed [33], a comparable matching has not been demonstrated in the case of a solid. In this case, the ideal reference model is given by the Einstein crystal, that is, a collection of pointlike, non-interacting particles restrained in specified positions in space by means of harmonic potentials. As the ideal gas, also for this model the partition function (hence the full thermodynamics) can be computed analytically: it thus represents the reference starting point of a nonperturbative computation of free energies and other quantities of interest.

The goal of this paper is to provide a detailed and comprehensive overview of the Hamiltonian adaptive resolution method H-AdResS [30,31,25] in the perspective of the coupling between a reference, high-resolution system and its ideal counterpart. Particular attention is given to the technical aspects of these simulations, with thorough descriptions of the computational and algorithmic characteristics. We will review the basics of the method and the satellite algorithms developed to improve its effectiveness, and discuss their application to two significant case studies: one is liquid water, described at the all-atom level, coupled to an ideal gas; the other one is a Lennard-Jones

solid coupled to an Einstein crystal. This second application is presented here for the first time, and constitutes the main original contribution of this work.

The manuscript is organised as follows: in sect. 2 we review the specific approach employed here, that is, the Hamiltonian Adaptive Resolution Simulation scheme, or `H-AdResS`; in sect. 3 we illustrate the significance of the external field required to obtain a uniform density profile, and describe in detail the algorithms developed to efficiently parametrise it; in sect. 4 we present the coupling between an atomistic model of water and an ideal gas, and describe the treatment of the electrostatic interaction within the `H-AdResS` framework; in sect. 5 we concentrate on the Lennard-Jones crystal and its coupling with the ideal Einstein crystal model; finally, in sect. 6 we summarise our results and provide a brief account of the possible applications of the presented methods.

2 Theoretical background

The `H-AdResS` scheme [30,31] belongs to a family of adaptive resolution simulation methods [34,22,23] in which a small portion of the system, usually a fluid described with atomistic resolution (AT), is embedded in a reservoir of particles of the same system, modelled using a coarse-grained (CG) representation; the coupling between the two resolutions takes place in an open boundary region, the so-called hybrid (HY) region. A snapshot of the simulation setup is presented in fig. 1. The main feature of the `H-AdResS` method is that the whole system is described in terms of a global Hamiltonian function H of the form

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

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

Several options are available to define the geometry of the all-atom region, *e.g.* spherical [26], cylindrical [35], or even with a time-dependent shape [36]. Here, without loss of generality, we employ one of the most common choices, namely a rectangular simulation box where the AT is represented by a slab. The resolution of a given molecule is thus determined by the following piecewise function:

$$\lambda(x) = \begin{cases} 1 & |x| \le d_{AT}/2 \\ \cos^2\left(\dfrac{\pi(x - d_{AT}/2)}{2 d_{HY}}\right) & \dfrac{d_{AT}}{2} < |x| \le \dfrac{d_{AT}}{2} + d_{HY} \\ 0 & |x| > d_{AT}/2 + d_{HY}, \end{cases}$$
$$(2)$$

with d_{AT} and d_{HY} the sizes of the AT and HY region, respectively. The mid-point of the simulation box is set in the origin of the coordinate system. Concerning non-bonded interactions, a molecule α interacts with its neighbouring particles through coarse-grained V^{CG} and atomistic V^{AT} potentials, defined as
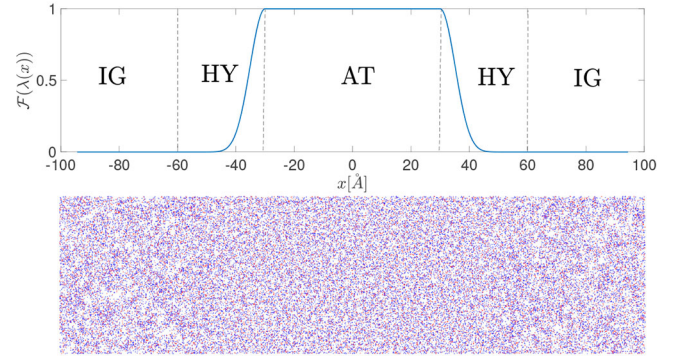

Fig. 1. Setup of a Hamiltonian adaptive resolution simulation. The periodic box is partitioned into three different regions, namely: ideal gas (IG), hybrid (HY), and atomistic (AT). Upper panel: the switching function $\mathcal{F}(\lambda) = \lambda^7$ takes values between 0 (IG) and 1 (AT), thus defining the resolution of a molecule (here water). Lower panel: simulation snapshot explicitly showing the various subdomains. The oxygen and hydrogen atoms of the water molecules are represented by red and blue dots, respectively. Note the different structure attained by the molecules in the different subdomains: a more uniform one in the AT region, and a more irregular one in the IG region. This difference is due to the ordering effect of the interactions among atomistic water molecules, imposing in particular the characteristic tetrahedral arrangement due to hydrogen bonds: this effect is naturally absent in the IG region, where molecules are randomly distributed.

$$V_{\alpha}^{AT} \equiv \frac{1}{2} \sum_{\beta, \beta \neq \alpha}^{N} \sum_{ij} V^{AT}(|\mathbf{r}_{\alpha i} - \mathbf{r}_{\beta j}|)$$

$$(3)$$

$$V_{\alpha}^{CG} \equiv \frac{1}{2} \sum_{\beta, \beta \neq \alpha}^{N} V^{CG}(|\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|),$$

where the coordinates of an atom —labeled by i, j indices— or a molecule —labeled by α, β indices— are represented by vectors $\mathbf{r}$ and $\mathbf{R}$, respectively. The factor $1/2$ accounts for the double counting of particles in the sum since the total non-bonded potential of a molecule α is given by a sum of AT and CG contributions weighted by λ_{α} or $(1 - \lambda_{\alpha})$, respectively. Albeit two-body interactions have been employed here, the extension to multibody, short-ranged potentials is straightforward. It is also worth pointing out that in this context the usage of the term *coarse grained* is referred to the particular type of interaction employed in the low-resolution region, *i.e.* an effective, simplified one with respect to the atomistic resolution model; the structure of the simulated molecules, on the other hand, is not affected by the change in resolution, *i.e.* they retain their fully atomistic structure throughout the simulation box. This is not a prerequisite to perform this kind of simulations, however it is a simple and effective manner to deal with the resolution change at the algorithmic and computational level in general, and is instrumen-

tal in the construction of a Hamiltonian-based approach in particular [30].

The total force acting on atom i of molecule α is given by

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

where $\mathbf{F}_{\alpha i|\beta}^{Res}$, $Res = \mathrm{AT, CG}$ is the force acting on atom i of molecule α due to its (atomistic or coarse-grained) interaction with the whole molecule β. The forces due to intramolecular interactions, $\mathbf{F}_{\alpha i}^{int}$, are not resolution dependent. The second term on the right-hand side 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. Each of the terms in the sum 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 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 $(V_\alpha^{AT} - V_\alpha^{CG})$. This *drift force* term $\mathbf{F}_{\alpha i}^{dr} = -[V_\alpha^{AT} - V_\alpha^{CG}] \nabla_{\alpha i} \lambda_\alpha$ violates Newton's Third Law and induces a pressure imbalance across the HY region, as it pushes molecules in the subdomain where the Helmholtz free energy is locally lower [30,31].

In general, the AT and CG representations of the same physical system follow different equations of state. Once coupled together *via* an open boundary, the AT and CG regions exchange particles to balance the differences in equilibrium pressure and chemical potential. This results, as it has already been thoroughly investigated [30,31,25], in a non-homogeneous density profile. To overcome these effects and enforce a uniform density profile, it is possible to introduce a new term in the Hamiltonian:

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

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 parametrise the term $\Delta H(\lambda)$.

3 Calculation of the free energy compensations

3.1 Compensation of the drift force

The free energy compensation term $\Delta H(\lambda)$ that neutralises, on average, the effect of the drift force, has to satisfy the relation:

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

in such a way that the total drift force resulting from the modified Hamiltonian reads

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

and, by construction, $\langle \hat{\mathbf{F}}_\alpha^{dr} \rangle = 0$.

Thermodynamic integration [19] has been originally proposed to compute the compensation term (6). This procedure can be justified on the basis that the potential $\Delta H(\lambda)$ bears a strict relationship with the Helmholtz free energy difference between a system with a Hamiltonian $H(\lambda)$ and a reference system defined by the CG representation at $\lambda = 0$ [30,25]. This procedure, apparently the most straightforward route to compute $\Delta H(\lambda)$, presents two main drawbacks: i) it requires an additional free energy calculation, and ii) the accuracy of the obtained $\Delta H(\lambda)$ can be limited when the system displays strong correlations within the hybrid region.

In a previous paper [37], some of us have shown that a more effective strategy is to compute and locally balance the drift force and parametrise the compensation *on the fly* by using an iterative scheme [25]. A given number N_b of bins is used to discretise the hybrid region in terms of λ in such a way that the width of the bins is $\delta\lambda = 1/N_b$. The index of the bins is given by $i = \mathtt{floor}[\lambda(\mathbf{R}_\alpha)/\delta\lambda]$. For a molecule α in a bin i, the contributions V_α^{CG} and V_α^{AT} are computed and accumulated in the local variables V_i^{CG} and V_i^{AT}, respectively. Moreover, the variables N_i^{CG} and N_i^{AT} are defined to monitor throughout the simulation the number of molecules in the bin. These quantities are computed in the same routine and simultaneously with the AT and CG forces.

For all molecules in the hybrid region, this procedure is carried out for a time interval of duration Δt and successive n iterations are performed. At the end of the n-th interval, average AT and CG potentials, $\bar{V}^R[i,n] = V_i^R/N_i^R$ with $R = \mathrm{AT, CG}$, are computed. Subsequently, the variables V_i^R, N_i^R are emptied and the average procedure continues. These average values calculated at the end of the n-th cycle are used to compute the running average $\mathcal{V}_{i,n}^R$, defined by

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

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

This running average, for $n > 0$, is used to compute the different components of the compensation force. Specifically, for t such that $t_0 + n\Delta t < t \leq t_0 + (n+1)\Delta t$, a

molecule α in the bin i of the HY region will feel compensation forces of the form

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

with $s = +1$ for $R = $ AT and $s = -1$ for $R = $ CG. At each time step, this force is distributed to the atoms in the molecule α with weights proportional to the atom/molecule mass ratio.

The update of the running average continues until the compensation forces converge to a steady value for every bin i. Finally, the update is interrupted and the resulting force compensation is integrated as a time-independent resolution-based force field, and the corresponding energy compensation $\Delta H(\lambda)$ can be easily computed.

The compensation of the drift force ensures that the hydrostatic pressure becomes uniform across the simulation domain [30,25]. The AT and CG representations equilibrate at this reference pressure according to the equation of state of the model. This equilibration implies that the densities in the two regions might differ, and an additional compensation must be applied to the system to ensure a flat density profile.

3.2 Compensation of the density imbalance

The uniform pressure enforced by the application of the compensation of the drift force does not necessarily guarantee a uniform density for the whole system. A density gradient might appear in the HY region as a result of the two models, AT and CG, following different equations of state. There is no general reason to require that the systems features a uniform density profile. Indeed, the only crucial aspect is that the thermodynamical state of the high-resolution region is under control and, specifically, consistent with the *desiderata* set by the user. Different systems and investigations require different conditions and properties, and a uniform density could be not among them.

In the particular case under examination, however, it is our aim to perform a smooth coupling between systems in the same thermodynamical state point as defined in the NVT (canonical) ensemble. Since not all thermodynamical properties and parameters can be simultaneously equated between the two subregions, the pressure acting on molecules in the AT domain will differ from that in the IG domain. The density, on the other hand, can be set to attain the same value if the compensation ΔH is modified in such a way that in each subregion the corresponding model attains a pressure that gives the same reference density.

This correction can be obtained via an iterative scheme dubbed *thermodynamic force* calculation [24], which consist in successively applying to the molecules in the HY region a force proportional to the density gradient:

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

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 convergence to a uniform density is guaranteed by the fact that the scheme has a fixed point when $\nabla\rho = 0$. In addition to ensuring the same density in the two subdomains, this method also leads to a flat density profile in the HY region [24,30,25].

Usually, 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. A new density profile is thus obtained and a new force is calculated. 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.

The iterative calculation of the thermodynamic force relies on an accurate estimation of the density profile, which 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 advantage of this procedure is that small deviations of the density from the reference value are immediately suppressed. However, it is evident that the density profile obtained from a small time interval ΔT is too noisy to reliably compute its numerical gradient.

To overcome this difficulty, the position of the centre of mass of the molecules is convoluted with a Gaussian function with a half-width $\sigma/2$, comparable with the typical excluded volume radius of the molecules. Thus, the density in the bin i, covering the coordinate range $[x_i, x_{i+1}]$ in a specific simulation frame, is computed as

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

The parameter l, whose appropriate value is system dependent, controls the range of the Gaussian function. A sensible choice is to set $l = 2.5\sigma$.

4 Water–ideal gas coupling

The reference, analytically solvable model of a liquid is the ideal gas (IG): in a computer simulation, this can be implemented as a collection of particles (with or without internal degrees of freedom) which do not interact. If subject only to the classical equations of motion and in absence of intermolecular potential, these particles would move along rectilinear trajectories with constant orientation and velocity (this is, assuming that the simulation box features periodic boundary conditions and not hard walls). To prevent this, and to be consistent with our goal of simulating a system in the canonical ensemble, a Langevin thermostat acts on the molecules in the ideal gas region, thereby

imposing on them a stochastic, diffusive dynamics. On the other hand, the conservative part of the Hamiltonian becomes

$$H = \mathcal{K} + V^{int} + \sum_{\alpha} \mathcal{F}(\lambda_\alpha) V_\alpha^{AT}, \qquad (12)$$

where only the AT part of the interaction is present. $\mathcal{K}$ is the kinetic energy term and V^{int} is the potential resulting from intramolecular interactions. Note that the resolution function λ has been replaced by a more general function of it, $\mathcal{F}(\lambda)$. In fact, since particles can get extremely close to each other in proximity of the HY/IG interface, a systematic sampling of huge potential energies might result in unphysical ensemble averages $\langle V^{AT} \rangle$ needed for the calculation of the compensations. To ensure a satisfactory statistical sampling, we modulate the resolution *via* $\mathcal{F}(\lambda) = \lambda^\nu$ with $\nu \geq 1$, which also takes values between 0 and 1 (see fig. 1). The value $\nu = 7$ is sufficient to smoothen out divergent interactions. Another advantage of using an exponent $\nu \geq 1$ is that the effective HY/IG interface moves closer to the AT domain. This results in a more stable and controlled thermodynamic force convergence. Furthermore, the AT potential is capped at a distance $\hat{r}$ to suppress large forces that might result from overlapping molecules. That is

$$V^{AT}(r) = \begin{cases} V^{AT}(\hat{r}) - \left. \dfrac{\partial V^{AT}}{\partial r} \right|_{r=\hat{r}} (r - \hat{r}), & r < \hat{r} \\ V^{AT}(r), & r \geq \hat{r} \end{cases} . \quad (13)$$

In the case of liquid water at room temperature, the overlapping events are rare (one in 0.5 ns, approximately). Moreover, they do not affect thermodynamic or structural properties of the system, as we have verified in the case of SPC/E water model. Furthermore, they do not contribute significantly to the calculation of the compensations since high-energy contributions, related to substantial overlaps of the molecules, are suppressed by the excluded volume and restricted to the tail of the configurational probability distribution.

The total force acting on the atom i of the molecule α (without the free energy compensation term) can be obtained from the Hamiltonian (12) as

$$\begin{aligned} \mathbf{F}_{\alpha i} = {}& \mathbf{F}_{\alpha i}^{int} \\ & + \sum_{\beta, \beta \neq \alpha} \left\{ \frac{\mathcal{F}(\lambda_\alpha) + \mathcal{F}(\lambda_\beta)}{2} \mathbf{F}_{\alpha i | \beta}^{AT} \right\} \\ & - V_\alpha^{AT} \left. \frac{\partial \mathcal{F}}{\partial \lambda} \right|_{\lambda=\lambda_\alpha} \nabla_{\alpha i} \lambda_\alpha. \end{aligned} \qquad (14)$$

Finally, concerning electrostatic interactions, we have used an alternative to the Ewald summation method, *i.e.* the damped shifted potential (DSF) [38,39]. In this approach, previously discussed in the context of H-AdResS simulations in [37], two charges q_i and q_j separated by a distance r_{ij} follow the electrostatic potential given by

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

where $r_{ij} \leq r_c$ and ϵ_0 is the vacuum permittivity. Only two parameters, the cut-off radius r_c and the damping parameter α, need to be specified. $\text{erfc}(r)$ is the complementary error function that takes into account the damping proposed in [38]. The gradient of the potential (15) gives the force acting on atom i

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

This framework is employed to perform MD simulations in the canonical ensemble for 15615 SPC/E [40–42] water molecules in a simulation box of size $188 \times 50 \times 50$ Å. The initial fully atomistic equilibrated configuration has been obtained from a simulation of 100 ps in the NPT ensemble followed by a 0.1 ns equilibration run performed with a time step of $\delta t = 0.001$ ps. The temperature and pressure are enforced at $T = 298$ K and $P = 1$ bar using the Nosé-Hoover thermostat and barostat with damping coefficient of 0.1 ps and 1 ps, respectively. The parameters for the DSF electrostatic potential, damping coefficient $\alpha = 0.2$ Å^{-1} and cut-off radius $r_c = 12$ Å, were chosen to reproduce the RDFs of reference simulations using the Ewald summation method. The Lennard-Jones and DSF potentials are capped at a radius $\hat{r} = 0.5$ Å.

Once the initial configuration has been obtained, the system density is set to the final value attained after the NPT equilibration run, and a H-AdResS simulation is performed using the Hamiltonian of eq. (12) without any compensation terms. After 100 ps, the *on the fly* calculation of the drift force compensation is applied, with updates every 5 ps, during 3000 ps. The resolution interval is divided into 1000 bins of size $\Delta\lambda = 0.001$. The *on the fly* density balancing method is applied simultaneously to the drift force correction. In this case, the length of the simulation box is uniformly discretised into slabs of size $\Delta x = 1.0$ Å and the thermodynamic force is updated every 50 ps. We employed values of $c = 1$ kcal mol^{-1} Å$^{-1} = 4.184$ kJ mol^{-1} Å^{-1}, $\sigma = 3$ Å and $l = 6$ Å, empirically optimised to maximise the stability of the algorithm, for smoothing and scaling the thermodynamic force. All simulations are performed with the LAMMPS simulation package [43,37], where the method is implemented, freely available and ready to use; the same algorithm is currently being implemented in the ESPResSO++ [44] software platform. The results presented in fig. 2 for the compensation terms obtained with this method show a rather smooth behaviour that validates the present approach.
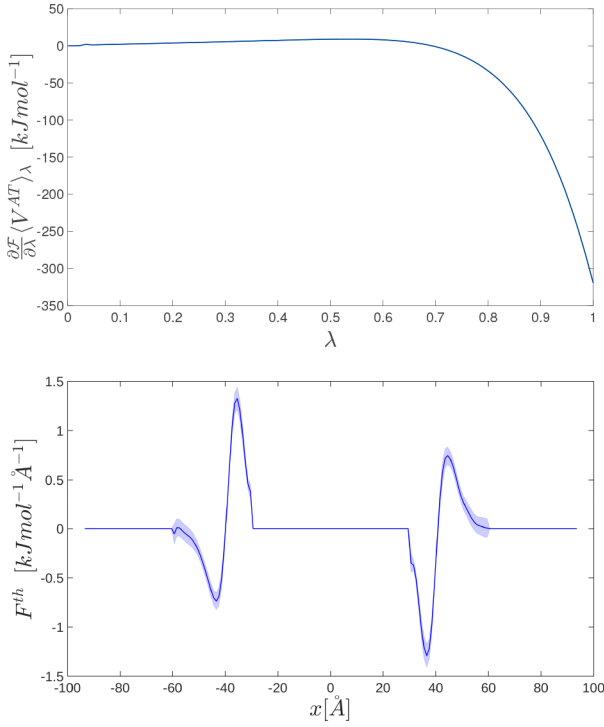
Fig. 2. Force compensations obtained from the procedure described in the text. Upper panel: compensation to the drift force as a function of λ. The use of the $\nu = 7$ coefficient guarantees that the contributions close to the IG/HY interface ($\lambda \sim 0$) are negligible compared to the contributions close to the AT region. Lower panel: thermodynamic force as a function of the position within the simulation box.

The successful coupling between a dense and strongly interacting fluid, such as water modelled at the atomistic level, and the ideal gas reference model has been previously demonstrated [45]: it is our scope, however, to include this verification in order to make the case for a sound and effective concurrent usage of atomistic and ideal models of different systems. The analysis of radial distribution functions (RDFs) and velocity autocorrelation functions (VACF) has been performed to check that the ideal gas reservoir bears no effect on structural, thermodynamical, and dynamical properties of the liquid in the all-atom subdomain. The results obtained in the dual-resolution setup are compared with the same quantities computed in fully atomistic benchmark simulation. In the case of the RDFs, fig. 3 shows a remarkable agreement between the two cases.

The velocity autocorrelation function (VACF) is defined as [5]:

$$C_{vv}(t) = \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle, \quad (17)$$

where $\mathbf{v}_i(t)$ is the velocity of molecule i at time t. To compute the VACF for the t_m-th time step the discrete estimator described in ref. [46] was used:

$$C_{vv}(t_m) = \frac{1}{N_{AA}} \sum_{i=1}^{N_{AA}} \frac{1}{M-m} \sum_{n=0}^{M-m-1} \mathbf{v}_i(t_{n+m}) \cdot \mathbf{v}_i(t_n), \quad (18)$$
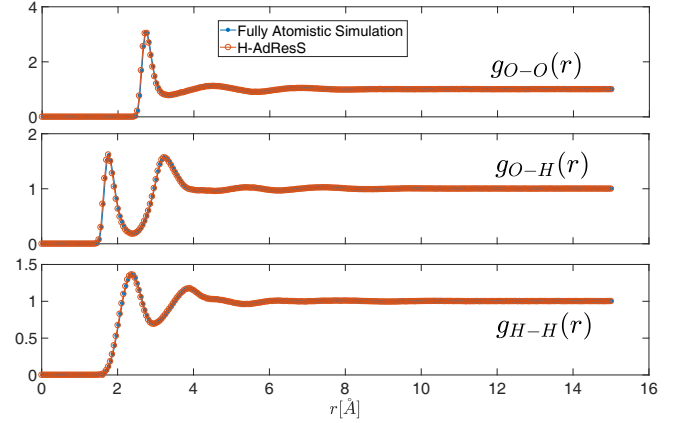

Fig. 3. RDFs of water molecules at pressure $P_0 = 1$ bar and temperature $T_0 = 298$ K in two different simulation setups: fully atomistic simulation (blue line with dots) and H-AdResS (red line with open circles). From top to bottom, the plots show oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen RDFs. The DSF damping parameter is set to $\alpha = 0.2$ Å^{-1}, and the cut-off radius is $R_c = 12$ Å.
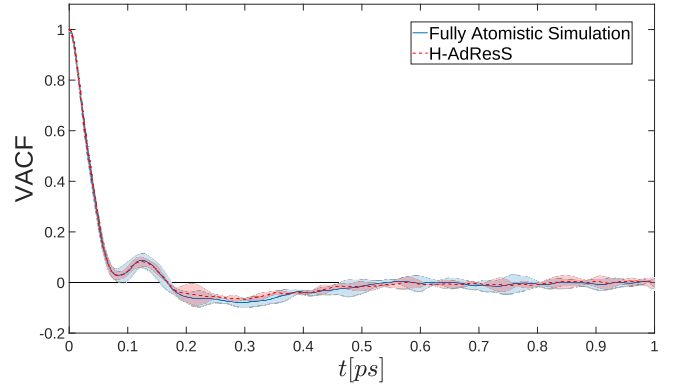

Fig. 4. Normalized velocity autocorrelation function ($C_{vv}(t)/C_{vv}(0)$) calculated for a reference fully atomistic (blue curve) and for the H-AdResS (red curve) simulations. In both cases only the atoms with coordinates in the interval -20 Å $< x < 20$ Å have been taken into account.

with M the total number of time steps and $t_m = m\delta t$, where δt is the integration time step. 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 $2t_{corr}/N_{AA}t_{tot}$ [5], with t_{corr} the correlation time and t_{tot} the total time of the simulation.

The results are reported in fig. 4. These show that a perfect consistency exists between the observables computed in the all-atom reference simulations and those obtained from the H-AdResS runs.

5 Solid–Einstein crystal coupling

The possibility of modulating the resolution of a system as a function of the position in space of its parts can be

easily extended from fluids to solids. Indeed, the field of material science has been the cradle of adaptive resolution simulation methods, since these were initially developed to perform *in silico* studies of *e.g.* crack propagation [47–52].

Here we describe the coupling of two different models of crystal within the framework of the H-AdResS method. Our objective, however, is not (only) to reduce the computational resources necessary to simulate large chunks of a solid of which only a tiny part, such as the expanding crack, is of interest; rather, we aim at two other goals: on the one hand, we want to describe how to practically perform an adaptive resolution simulation of a solid using, as a coarse-grained model, an extremely simple and exactly solvable representation, expanding the atomitstic-to-ideal gas coupling outside of the realm of liquids. On the other hand, we want to demonstrate that a substantial advantage can come from this coupling, namely the possibility to extract important thermodynamical information on the system's free energy.

Arguably, the most idealised particle-based representation of a solid is the ideal Einstein crystal (IEC) where N non-interacting particles are coupled to their lattice sites using harmonic potentials. The system's potential energy has the form

$$V^E = \sum_{i=1}^{N} V_i^E = \frac{1}{2} \sum_{i=1}^{N} \kappa_i (\mathbf{r}_i - \mathbf{r}_i^0)^2, \qquad (19)$$

where each particle i fluctuates around its equilibrium position $\mathbf{r}_i^0$ with spring constant κ_i. To illustrate the method, as high-resolution model we employ a Lennard-Jones (LJ) potential V^{LJ}, whose well depth ϵ and excluded volume size σ set the energy and length scale, respectively. The corresponding H-ADResS Hamiltonian has the form

$$H = \mathcal{K} + \sum_{i=1}^{N} \left\{ \lambda_i V_i^{LJ} + (1 - \lambda_i) V_i^E \right\}, \qquad (20)$$

with $\mathcal{K}$ being the kinetic energy term and N the total number of particles in the system. This Hamiltonian generates the following total force acting on the atom i:

$$\mathbf{F}_i = \sum_{j \neq i}^{N} \left(\frac{\lambda_i + \lambda_j}{2} \mathbf{F}_{ij}^{LJ} \right) + (1 - \lambda_i) \mathbf{F}_i^E$$
$$- \left[V_i^{LJ} - V_i^E \right] \nabla_i \lambda_i. \qquad (21)$$

Since, by construction, the density is uniform throughout the simulation box, the only energy compensation needed is the one required to counteract the drift force, *i.e.*, the Helmholtz free energy difference between the LJ and the IEC models. Therefore, the Hamiltonian in eq. (20) becomes

$$H_\Delta = H - \sum_{i=1}^{N} \Delta H(\lambda(\mathbf{r}_i)), \qquad (22)$$

with

$$\left. \frac{\mathrm{d}\Delta H(\lambda(\mathbf{r}_i))}{\mathrm{d}\lambda(\mathbf{r}_i)} \right|_{\lambda = \lambda_i} = \left\langle \left[V_i^{LJ} - V_i^E \right] \right\rangle_{\mathbf{r}_i}, \qquad (23)$$
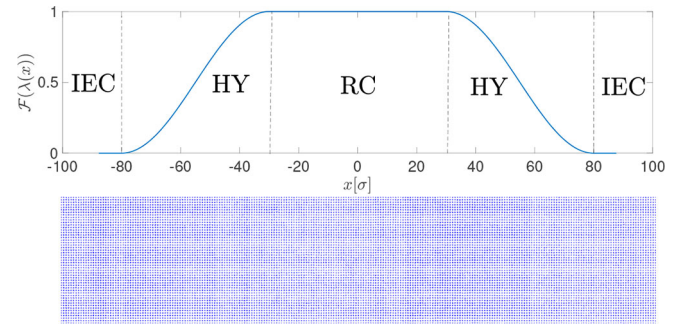

Fig. 5. Setup of a Hamiltonian adaptive resolution simulation for solids. The periodic box is divided into three different regions, namely: ideal Einstein crystal (IEC), hybrid (HY), and real crystal (RC). Upper panel: the switching function $\mathcal{F}(\lambda)$ takes values between 0 (IG) and 1 (AT), thus defining the resolution of a molecule. Lower panel: simulation snapshot explicitly showing the various subdomains. The system is composed of Lennard Jones particles at number density $\rho = 1.28\sigma^{-3}$ and the thermal energy $k_B T = 2.0\epsilon$.

Analogously to the case of fluids, we can integrate the previous expression and obtain the Helmholtz free energy compensation:

$$\Delta H(\lambda(\mathbf{r}_i)) = \int_0^\lambda \mathrm{d}\lambda' \left\langle \left[V_i^{LJ} - V_i^E \right] \right\rangle_{\mathbf{r}_i}. \qquad (24)$$

In the upper panel of fig. 5 we provide a schematic representation of the simulation box and the different subregions: real crystal (RC), hybrid (HY), and Einstein crystal (IEC), together with $\lambda(x)$. The lower panel of the same figure shows a snapshot of the simulation setup.

Equation (24) is analogous to the difference in Helmholtz free energy as obtained from the thermodynamic integration (TI) method proposed by Frenkel and Ladd [53–56]. The free energy of an IEC can be computed analytically, hence it can be used as the reference to calculate the Helmholtz free energy of a target crystal through a regular TI. To perform the latter, in our notation we write a Hamiltonian of the form

$$H(\lambda) = \mathcal{K} + \lambda V^{LJ} + (1 - \lambda) V^E, \qquad (25)$$

where in this case λ is a *global* coupling constant. The derivative of the free energy F with respect to λ gives (with $\beta^{-1} = k_B T$):

$$\frac{\partial F}{\partial \lambda} = -\beta^{-1} \frac{\partial}{\partial \lambda} \left[\ln \int \mathrm{d}\mathbf{r}^N \exp(-\beta H(\lambda)) \right]$$
$$= \left\langle V^{LJ} - V^E \right\rangle_\lambda, \qquad (26)$$

and the free energy of a real crystal F^{RC} is related to the free energy of the Einstein crystal F^{IEC} by

$$F(\lambda = 1) = F(\lambda = 0) + \int_0^1 \mathrm{d}\lambda \left\langle V^{LJ} - V^E \right\rangle_\lambda, \qquad (27)$$

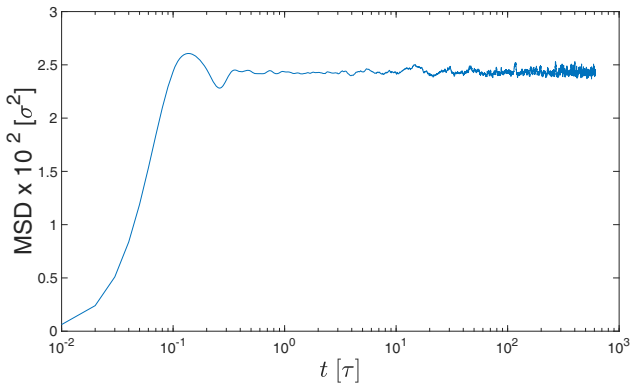with $F(\lambda = 1) = F^{RC}$ and $F(\lambda = 0) = F^{IEC}$.

Fig. 6. Mean squared displacement (MSD) of particles obtained from a reference LJ simulation. The curve reaches a plateau after 1 τ, and the asymptote is 0.024 σ^2.

Fig. 7. Atomistic per particle potential V^{AT} (upper panel) and coarse-grained per particle potential V^{CG} (lower panel) as a function of λ. In both panels, diamond, circle, and square data points are obtained from H-AdResS, thermodynamic integration, and fixed centre-of-mass (CoM) thermodynamic integration, respectively.

To compare the results obtained employing the two methods, it is necessary to set the value of the spring constant κ in eq. (19). A criterion to fix this parameter is to ensure that the mean squared displacement (MSD) of the particles in the fully atomistic simulation equals that of the ideal Einstein crystal, *i.e.* [53,54]:

$$\frac{3}{\kappa\beta} = \left\langle \left(\mathbf{r}_i - \mathbf{r}_i^0\right)^2 \right\rangle, \qquad (28)$$

For all cases, a simulation box of size $188.8\sigma \times 49.96\sigma \times 49.96\sigma$ is used. Initially, the LJ particles are placed on the fcc lattice structure with number density $\rho = 1.28\sigma^{-3}$ and their initial positions are set as the equilibrium positions (see eq. (19)). In H-AdResS, the sizes of the LJ crystal and hybrid regions are set to be 60σ and 50σ respectively (see fig. 5). The temperature $k_B T = 2\epsilon$ is fixed by a Langevin thermostat with damping coefficient of 10τ and the total net force acting on the particles is periodically removed. The particles are thermalised in the hybrid and IEC regions only, yet the temperature is uniform throughout the system (data not shown). In all simulations, the LJ potential is truncated at a cut-off radius $R_c = 2.7\sigma$ and not shifted [56]. The LJ potential is capped at a radius $\hat{r} = 0.1\sigma$. The time step is $\delta t = 0.001\tau$. To obtain the drift force, the resolution interval is divided into 20 bins of size $\Delta\lambda = 0.05$, and every 50000 time steps the *on the fly* calculation of the drift force compensation is performed and then updated. The duration of each simulation run is at least 2×10^6 steps.

The time evolution of the MSD of the Lennard-Jones crystal is presented in fig. 6. The curve shows the characteristic behaviour for a solid, where an initial superdiffusive regime is followed by a plateau, which in our case appears after 1 τ. From the asymptote at 0.024 σ^2 we obtain the spring stiffness $\kappa = 125k_B T$. However in both the H-AdResS and the TI simulations we employed $\kappa = 250k_B T$, as this value guarantees the potential to be stiff enough to avoid neighbouring particles to collapse, and flexible enough to properly sample the interaction.

To validate the consistency of the method, we computed the two components of the derivative of the per particle Helmholtz free energy, namely $\langle V^{LJ} \rangle$ and $\langle V^{IEC} \rangle$,
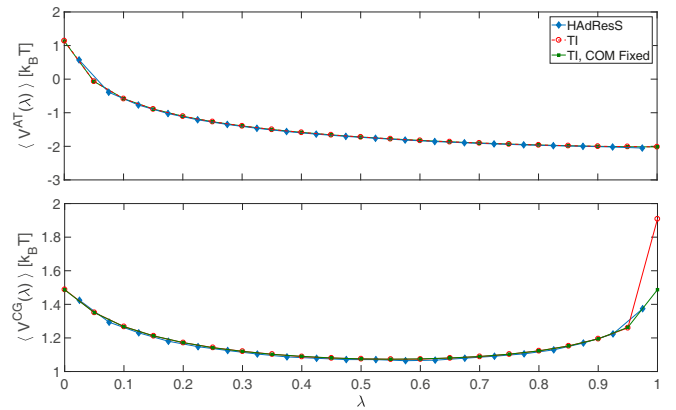
making use of the H-AdResS method and the TI, with and without the fixed centre of mass (CoM) constraint. These data are reported and compared in fig. 7: the three data sets agree remarkably well, with the only exception of the point $\lambda = 1$ for the regular TI. This is expected, because in the TI without restraints the LJ crystal is kept at a fixed position in space only through the coupling with the IEC. When $\lambda \to 1$ this constraint relaxes, and the LJ crystal gets asymptotically free to diffuse away from the IEC. For this reason, it is customary to include in the TI simulation a constraint to the LJ crystal centre of mass position. Conversely, in the H-AdResS simulation the LJ crystal cannot drift because of the explicit and permanent spatial coupling to the IEC. This is made evident by the excellent agreement for the full range of λ between the H-AdResS and the data obtained from a TI integration with fixed CoM. More generally, we expect the averages obtained in the adaptive resolution simulation to correspond point by point to the ones from TI. This expectation stems from the fact that, for sufficiently large hybrid regions, the system is locally —*i.e.* for a given value of λ— in a state comparable to that of a TI setup with the same value of the Hamiltonian interpolation parameter [25]. Furthermore, we know from previous work [30, 31,25] that the free energy compensation corresponds to the Helmholtz free energy difference between the all-atom and the coarse-grained representation; since the difference between the integral of these quantities has to sum up to the same value for both TI-based and H-AdResS-based calculations, and excluding the possibility of exact cancellations between deviations present in different parts of the curves, we conclude that the integrands have to be point-wise identical. Indeed, the LJ crystal excess free energies $(F(\lambda = 1) - F(\lambda = 0))$ computed with respect to the IEC with the two approaches discussed here are in excellent agreement: the reference calculation by means of standard TI with constraints gives $-2.63k_B T$ while the H-AdResS-based approach provides a value of $-2.64k_B T$.
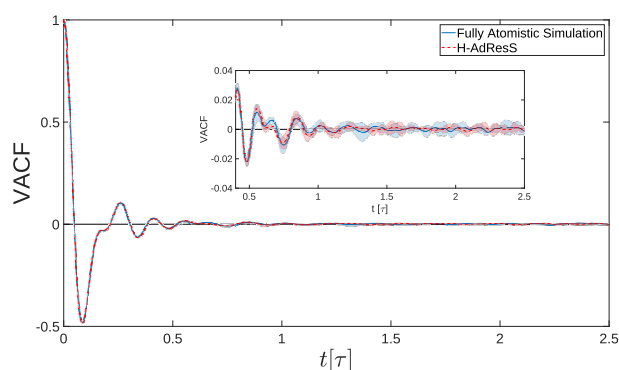
Fig. 8. Normalized velocity auto-correlation function $(C_{vv}(t)/C_{vv}(0))$ of LJ solid. The result of the fully atomistic, reference simulation is shown with the solid blue curve. The result of the H-AdResS simulation is shown in red dashed curve. The simulation box has dimensions of $188.8\sigma \times 49.96\sigma \times 49.96\sigma$. In both cases, only those atoms with coordinates such that $-20\,\sigma < X < 20\,\sigma$ are considered for the calculation of the VACF.

As for the case of liquid–ideal gas coupling, also here we can monitor the consistency of the equilibrium dynamical properties of the crystal in the high-resolution region of the H-AdResS setup. This is done by calculating the VACF for the fully atomistic and the dual-resolution setup, whose results are presented in fig. 8. In both cases the VACF shows the same behaviour within error bars, the latter being indicated by the shaded regions around the solid lines. As expected for a solid, the curves fluctuate around zero with amplitudes decreasing with time. These fluctuations result from the incoherent vibrations of the particles at their equilibrium positions. This is a remarkable result, since for an ideal Einstein crystal the VACF exhibits a sinusoidal behaviour: in the dual-resolution simulation, however, the dynamics of the LJ subdomain are not affected by the harmonic character of the IEC.

6 Conclusions and outlook

Ideal models, in the context of soft matter, are characterised by two elements: absence of interactions (or at least coupling) and exact integrability. This makes them historically pivotal in the comprehension of many physical systems, as they entail a great fraction of the complexity of the latter in spite of a bare-to-the-bones nature. Semi-perturbative corrections can take quite far from the initial simplicity and a lot closer to physical reality, as it is the case of van der Waals' equation of state for real liquids. When such a simple modification, which still leaves the integrability of the model substantially intact, is not possible, other routes can open up, such as the decomposition of physical quantities in ideal and excess terms: the problem is now flipped from a perturbative modification of the ideal case to a finite, often substantial yet additive contribution, whose calculation is made possible by analytical or computational tools.

Adaptive resolution simulations methods are the most suitable instruments to take advantage of the comparison between a realistic description of a system and the corresponding ideal models: this is because such comparison is as direct as possible, with the two models coexisting in the same simulation and in direct contact, exchanging particles and energy. To enforce the equality of a given property in both subdomains at different resolution implies to compute the differences that exist between them, often in a simpler and more effective way than by means of well established techniques. In this work we have revised the application of the Hamiltonian adaptive resolution simulation scheme H-AdResS to the coupling of all-atom models of a liquid to an ideal gas, and have demonstrated how an analogous setup can be constructed for a solid. The subdomain where the system is described with an interacting, accurate model does not suffer from the coupling, as it is shown by all structural and dynamical quantities measured there, whose values are in excellent agreement with the fully high-resolution reference. For the solid this is particularly relevant and remarkable, because the particle arrangement is not, as in a liquid, transitional and short-lived, rather remains fixed. Hence, any difference in thermodynamical and equilibrium dynamical properties with respect to the LJ solid could penetrate more strongly into the high-resolution domain, without being averaged out by fluctuations. Nonetheless, the switch from the interacting system to the ideal one, that takes place in the hybrid region, is smooth enough as to decouple the fine details of each model's properties, yet preserving the same thermodynamics throughout the simulation box.

These results provide the conceptual and technical bedrock for applications with important consequences. The computational cost of these simulations, in fact, is reduced to that of the high-resolution domain, as the ideal gas/solid parts require a negligible amount of force calculations. The size of systems that can be treated with these technique can thus be substantially large, so as to minimise finite-size effects and enable the simulation of complex molecules. As a particularly useful perspective application we envisage the calculation of chemical potentials of complex molecular crystals, a topic of increasing relevance for its industrial and pharmaceutical implication, which is the object of ongoing work.

Author contribution statement

MH, RCH, KK and RP conceived the problem. MH performed the simulations. MH, RCH and RP analysed the data. MH, RCH, KK and RP wrote the paper.

References

1. R.P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
2. D. Frenkel, J.-P. Hansen, Phys. World **9**, 35 (1996).
3. W.F. van Gunsteren, A.E. Mark, J. Chem. Phys. **108**, 6109 (1998).
4. W.G. Hoover, *50 Years of Computer Simulation – a Personal View*, arXiv:0812.2086v2 (2008).
5. M.P. Allen, D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).
6. D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier, 2001).
7. M. Praprotnik, L. Delle Site, K. Kremer, Annu. Rev. Phys. Chem. **59**, 545 (2008).
8. K. Kremer, *Soft and fragile matter non equilibrium dynamics, metastability and flow*, in *SUSSP Proceedings* Vol. **53** (IOP Publishing Ltd., 2000) pp. 145–184.
9. A. Mulero (Editor), *Theory and Simulation of Hard-Sphere Fluids and Related Systems* (Springer, Berlin, Heidelberg, 2008).
10. K. Kremer, F. Müller-Plathe, MRS Bull. **26**, 205 (2001).
11. R.E. Caflisch, G. Ceder, K. Kremer, T. Pollock, M. Scheffler, E.G. Wang (Editors), *Focus on Novel Materials Discovery*, New J. Phys. (IOP, 2013 and 2014).
12. C. Micheletti, P. Carloni, A. Maritan, Proteins **55**, 635 (2004).
13. W.G. Noid, J. Chem. Phys. **139**, 090901 (2013).
14. F. Pontiggia, A. Zen, C. Micheletti, Biophys. J. **95**, 5901 (2008).
15. M. Karplus, J. McCammon, Nature **277**, 578 (1979).
16. M. Karplus, Acc. Chem. Res. **35**, 321 (2002).
17. A. Pérez, F.J. Luque, M. Orozco, Acc. Chem. Res. **45**, 196 (2012).
18. P. Ballone, Entropy **16**, 322 (2014).
19. J. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
20. P. Raiteri, A. Laio, F.L. Gervasio, C. Micheletti, M. Parrinello, J. Phys. Chem. B **110**, 3533 (2006).
21. M.E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, 2010).
22. M. Praprotnik, L. Delle Site, K. Kremer, Phys. Rev. E **73**, 066701 (2006).
23. M. Praprotnik, L. Delle Site, K. Kremer, J. Chem. Phys. **126**, 134902 (2007).
24. S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, K. Kremer, Phys. Rev. Lett. **108**, 170602 (2012).
25. P. Español, R. Delgado-Buscalioni, R. Everaers, R. Potestio, D. Donadio, K. Kremer, J. Chem. Phys. **142**, 064115 (2015).
26. A.C. Fogarty, R. Potestio, K. Kremer, J. Chem. Phys. **142**, 195101 (2015).
27. A.C. Fogarty, R. Potestio, K. Kremer, Proteins **84**, 1902 (2016).
28. R. Fiorentini, K. Kremer, R. Potestio, A.C. Fogarty, J. Chem. Phys. **146**, 244113 (2017).
29. T. Tarenzi, V. Calandrini, R. Potestio, A. Giorgetti, P. Carloni, J. Chem. Theory Comput. **13**, 5647 (2017).
30. R. Potestio, S. Fritsch, P. Español, R. Delgado-Buscalioni, K. Kremer, R. Everaers, D. Donadio, Phys. Rev. Lett. **110**, 108301 (2013).
31. R. Potestio, P. Español, R. Delgado-Buscalioni, R. Everaers, K. Kremer, D. Donadio, Phys. Rev. Lett. **111**, 060601 (2013).
32. M. Heidari, K. Kremer, R. Cortes-Huerto, R. Potestio, *Spatially resolved thermodynamic integration: An efficient method to compute chemical potentials of dense fluids*, arXiv:1802.08045, submitted to J. Chem. Theory Comput.
33. K. Kreis, A.C. Fogarty, K. Kremer, R. Potestio, Eur. Phys. J. ST **224**, 2289 (2015).
34. M. Praprotnik, L. Delle Site, K. Kremer, J. Chem. Phys. **123**, 224106 (2005).
35. J. Zavadlav, R. Podgornik, M. Melo, S. Marrink, M. Praprotnik, Eur. Phys. J. ST **225**, 1595 (2016).
36. K. Kreis, R. Potestio, K. Kremer, A.C. Fogarty, J. Chem. Theory Comput. **12**, 4067 (2016).
37. M. Heidari, R. Cortes-Huerto, D. Donadio, R. Potestio, Eur. Phys. J. ST **225**, 1505 (2016).
38. D. Wolf, P. Keblinski, S.R. Phillpot, J. Eggebrecht, J. Chem. Phys. **110**, 8254 (1999).
39. C.J. Fennell, J.D. Gezelter, J. Chem. Phys. **124**, 234104 (2006).
40. H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
41. L.X. Dang, B.M. Pettitt, J. Phys. Chem. **91**, 3349 (1987).
42. Y. Wu, H.L. Tepper, G.A. Voth, J. Chem. Phys. **124**, 024503 (2006).
43. S. Plimpton, J. Comput. Phys. **117**, 1 (1995).
44. J.D. Halverson, T. Brandes, O. Lenz, A. Arnold, S. Bevc, V. Starchenko, K. Kremer, T. Stuehn, D. Reith, Comput. Phys. Commun. **184**, 1129 (2013).
45. K. Kreis, A.C. Fogarty, K. Kremer, R. Potestio, Eur. Phys. J. ST **224**, 2289 (2015).
46. J. Kohanoff, Comput. Mater. Sci. **2**, 221 (1994).
47. F. Pavia, W.A. Curtin, Model. Simul. Mater. Sci. Eng. **23**, 055002 (2015).
48. R. Rudd, J. Broughton, Phys. Status Solidi B: Basic Res. **217**, 251 (2000).
49. J. Rottler, S. Barsky, M.O. Robbins, Phys. Rev. Lett. **89**, 148304 (2002).
50. G. Csanyi, T. Albaret, M.C. Payne, A.D. Vita, Phys. Rev. Lett. **93**, 175503 (2004).
51. D. Jiang, E.A. Carter, Acta Mater. **52**, 4801 (2004).
52. G. Lu, E.B. Tadmor, E. Kaxiras, Phys. Rev. B **73**, 024108 (2006).
53. D. Frenkel, A.J.C. Ladd, J. Chem. Phys. **81**, 3188 (1984).
54. J.M. Polson, E. Trizac, S. Pronk, D. Frenkel, J. Chem. Phys. **112**, 5339 (2000).
55. M.A. van der Hoef, J. Chem. Phys. **113**, 8142 (2000).
56. C. Vega, E.G. Noya, J. Chem. Phys. **127**, 154113 (2007).