Adaptive molecular resolution approach in Hamiltonian form: An asymptotic analysis

Jinglong Zhu
Institute for Mathematics, Freie Universität Berlin, D-14195 Berlin, Germany
and School of Mathematical Sciences, Peking University, Beijing 100871, People’s Republic of China

Rupert Klein and Luigi Delle Site*
Institute for Mathematics, Freie Universität Berlin, D-14195 Berlin, Germany

(Received 22 June 2016; revised manuscript received 23 September 2016; published 28 October 2016)

Adaptive molecular resolution approaches in molecular dynamics are becoming relevant tools for the analysis of molecular liquids characterized by the interplay of different physical scales. The essential difference among these methods is in the way the change of molecular resolution is made in a buffer (transition) region. In particular, a central question concerns the possibility of the existence of a global Hamiltonian which, by describing the change of resolution, is at the same time physically consistent, mathematically well defined, and numerically accurate. In this paper we present an asymptotic analysis of the adaptive process complemented by numerical results and show that under certain mathematical conditions a Hamiltonian, which is physically consistent and numerically accurate, may exist. Such conditions show that molecular simulations in the current computational implementation require systems of large size, and thus a Hamiltonian approach such as the one proposed, at this stage, would not be practical from the numerical point of view. However, the Hamiltonian proposed provides the basis for a simplification and generalization of the numerical implementation of adaptive resolution algorithms to other molecular dynamics codes.

DOI: 10.1103/PhysRevE.94.043321

I. INTRODUCTION

Large systems of molecular liquids are characterized by processes occurring at different scales which in turn often require a different level of accuracy regarding the molecular model [1]. Highly accurate molecular models lead to a complete physical picture but require large computational resources and additional work of analysis of the large amount of data produced; in fact a clear description of a process requires a screening of data to the essential. On the other hand, less accurate molecular models are computationally convenient and produce a small amount of data to analyze; however, they are likely to miss, due to their simplification, essential physical features [2]. In this perspective, multiscale methods in molecular simulation (MS) have been developed in the past few years for optimizing the need for a consistent physical treatment and acceptable numerical and analysis costs [1,2]. In particular, adaptive molecular resolution methods, which, by partitioning the system in regions of different molecular resolution, change the number of degrees of freedom on the fly, became very appealing due to successful numerical performance [3,4]. Various implementations of this method exist and they differ for the definition of the coupling between different regions [3–6]. In this work we propose a procedure for coupling the atomicistic and coarse-graining region via an interface region which acts as a filter to transform atomicistic resolution into coarse-graining resolution and vice versa. The coupling is done by considering the Hamiltonian of the interface as a perturbation to an otherwise exact Hamiltonian written as a sum of a full atomicistic and a full coarse-grained Hamiltonian. Next we perform a mathematical treatment in terms of asymptotic analysis involving characteristic lengths of the system. In this way we derive mathematical conditions which, if reasonably fulfilled in the numerical simulation, ensure that the system behaves practically as a Hamiltonian system. The existence of a global Hamiltonian, although not necessary for adaptive resolution simulations [6,7], could provide technical advantages in the implementation of the code. In fact all standard molecular dynamics codes are based on Hamiltonian algorithms and thus one may use their computational architecture in an almost straightforward way. In general the direct coupling of an atomistic system with a coarse-graining system is expected to lead to dissipation and thus different correction terms are added to take care of such a problem [4–9]. In particular in Hamiltonian-based algorithms the Hamiltonian is corrected by adding a free-energy term. Our aim instead is to provide a definition of a global Hamiltonian based solely on particles’ degrees of freedom; it is our opinion that a formula that contains free-energy terms does not define a proper Hamiltonian. If such a definition is possible, then the system would be self-contained in the sense that, once the simulation setup is defined, the numerical calculation can run without any need for further additional quantity calculated in additional simulations. In particular our partitioning of the system paves the way for a generalization of the algorithm to different codes. This can be done by a simplification of the computational algorithm offered by the partitioning we propose: the simulation (given the interaction cutoff) at a certain stage can be performed in two distinct regions (atomistic and coarse grained) and then a third region where they overlap (the region of perturbation). For each time step, force calculations (i.e., the most expensive part of the code) can effectively run in parallel in each of the two regions and then can be synchronized through the force calculation in the overlapping region. Work along this direction is currently in progress. Finally, it must be also clarified that all the adaptive methods cited before [4–7] from the mere numerical

*luigi.dellesite@fu-berlin.de
point of view are essentially equivalent and/or equivalently efficient. The conceptual difference we discuss here regards the formal background on which they are based and the physical interpretation of the global Hamiltonian in terms of statistical mechanics of their results (see also note 4 in Ref. [6]).

II. ADAPTIVE MOLECULAR RESOLUTION: FORCE BASED OR HAMILTONIAN BASED

In this section we review the basic principles employed in the construction of adaptive resolution schemes. The schemes can be classified in two major categories: (a) force-based interpolation schemes [2,3] and (b) potential-based (Hamiltonian-based) schemes [4,5,10–13].

A. Force-based and grand-canonical-like scheme

The adaptive resolution simulation (AdResS) method has been developed following a simple intuitive principle. Such a principle consists of dividing the space in three distinct regions: (i) an atomistic region (high resolution), (ii) a coarse-grained region (low resolution), and (iii) an interface or hybrid transition region where molecules change their resolution. Next the intuitive physical requirement is that the molecules of the atomistic region follow the rules of a standard atomicistic dynamics, the molecules of the coarse-grained region follow the rules of a standard coarse-grained dynamics, and when a molecule transits in the interface region its dynamics slowly passes from an atomistic type to a coarse-grained type (or vice versa). The meaning of “slowly” is that the perturbation due to the change of resolution to the dynamics of the atomistic and of the coarse-grained region is negligible in the calculations of physical quantities of interest. In molecular dynamics (MD) such a principle can be easily implemented by smoothly interpolating in space the atomistic and coarse-grained forces:

\[
F_{i,j} = w(r_{ij})F_{\text{AT}}^{i,j} + [1 - w(r_{ij})]F_{\text{CG}}^{i,j},
\]

where \(i\) and \(j\) are the indices of two distinct molecules, \(F_{\text{AT}}^{i,j}\) is the force derived from the atomistic potential \((U_{\text{AT}})\) and \(F_{\text{CG}}^{i,j}\) is the force derived from the coarse-grained potential \((U_{\text{CG}})\) (usually a c.m.-c.m. potential, where c.m. indicates the center of mass), \(r\) is the c.m. position of the molecule, and \(w(x)\) is a smooth function varying from 0 to 1 in the transition region (HY):

\[
w(x) = \begin{cases} 
1, & x < x_{\text{AT}} \\
\cos^2 \left( \frac{\pi}{2d} (x - x_{\text{AT}}) \right), & x_{\text{AT}} < x < x_{\text{AT}} + d \\
0, & x_{\text{AT}} + d < x,
\end{cases}
\]

with \(x_{\text{AT}}\) the \(x\) location of the border of the AT region (see Fig. 1).

In an effective way, the atomistic degrees of freedom are slowly removed when a molecule leaves the atomistic region and enters the coarse-grained region and vice versa. As it is described above, the scheme is dissipative, that is, the change in number of degrees of freedom implies a gain or loss of kinetic and potential energy which is not spontaneously balanced, and in fact the system, prepared following this scheme, is not Hamiltonian [14]. However, the coupling to an external thermostat, which takes care of adsorbing the excess or adding the missing energy, was proven to be technically sufficient for the scheme to work properly [15,16]. At a later stage this empirical finding was justified on the basis of first principles of thermodynamics and statistical mechanics and recently the basic scheme reported above was embedded into the formalism of a grand-canonical ensemble where the coarse-grained and transition regions act as a particle reservoir for the atomistic region [6–9,17–19]. The technical details are not relevant for this paper, except that in a grand-canonical-like setup, the reservoir acts in a stochastic way and thus the only Hamiltonian of relevance is that of the atomistic region which implies that the nonexistence of a global Hamiltonian is no more a problem. However, some branches of research followed a different path based on the search of a global Hamiltonian without the addition of any stochastic aspect. An overview of the progress and pitfalls of such approaches is reported below.

B. Hamiltonian-based schemes

In a series of papers Ensing and collaborators [4,10–12] presented a scheme based on the interpolation of potentials instead of forces as done in AdResS:

\[
U_{ij}^{\text{tot}} = s(r_{ij})U_{ij}^{\text{AT}} + [1 - s(r_{ij})]U_{ij}^{\text{CG}},
\]

where \(U_{ij}\) is the global potential of interaction between molecule \(i\) and molecule \(j\), \(s(r_{ij})\) is the smooth interpolation function (slightly but not substantially different from the function \(w\) used in AdResS), and \(U_{ij}^{\text{AT}}\) and \(U_{ij}^{\text{CG}}\) are the atomistic and coarse-grained potentials acting between molecule \(i\) and molecule \(j\). If from \(U_{ij}^{\text{tot}}\) one derives the force between molecule \(i\) and molecule \(j\) then the result is a force as that of AdResS plus an additional term: \(\nabla s(r_{ij})U_{ij}^{\text{AT}} - U_{ij}^{\text{CG}}\). In their work the dissipative action of this spurious force needs to be balanced in such a way that it disappears from the simulation statistics. To achieve this, the method uses book-keeping of the gain or loss of energy of molecules that change resolution and thus adds at each instant the gain or loss of energy due to the change of resolution and preserves a global Hamiltonian approach.

The accurate determination of the gain or loss of energy per molecule and its adsorption or release in the system was done by violating the request that the system should be self-contained. In fact, in essence, the system was coupled to a generic external thermostat and thus this method became...
equivalent to the force-based AdResS scheme. It was then shown that without the thermostat the system was highly dissipative and that the interpolation of potentials plus the book-keeping energy was not sufficient to build a conservative scheme [14]. However, later on, the idea of providing or removing energy without using a thermostat was technically implemented by a method named H-AdResS developed by Potestio and collaborators [5,13,20,21]. In H-AdResS the interpolation is also done as in Eq. (3) and it was observed that such an interpolation is equivalent to the scheme of thermodynamic integration for the calculation of the difference of free energy between a state A and a state B of the system. In this case state A and state B meant the passage from the atomistic potential to the coarse-grained potential and thus the interpolation formula is equivalent to the space-dependent atomistic potential to the coarse-grained potential and thus thermodynamic integration for the calculation of the difference of free energy between a state A and a state B of the system. Consider a typical adaptive resolution setup as illustrated in Fig. 1 and let us decompose the hypothetical global Hamiltonian of the system as

\[
H_{\text{glob}} = H_{\text{AT-AT}} + H_{\text{CG-CG}} + H_{\text{HY-HY}} + H_{\text{AT-HY}} + H_{\text{CG-HY}} + H_{\text{intra}},
\]

where

\[
H_{\text{AT-AT}} = \sum_{a=1}^{N_{\text{AT}}} \left( \sum_{i=1}^{s_a} \frac{p_{ia}^2}{2m_{ia}} \right) + \sum_{a=1}^{N_{\text{AT}}} \sum_{b=1, b \neq a}^{N_{\text{AT}}} V_{\text{AT}}(r_{ia}, r_{ib}),
\]

\[
H_{\text{CG-CG}} = \sum_{a=1}^{N_{\text{CG}}} \left( \sum_{i=1}^{s_a} \frac{p_{ia}^2}{2m_{ia}} \right) + \sum_{a=1}^{N_{\text{CG}}} \sum_{b=1, b \neq a}^{N_{\text{CG}}} V_{\text{CG}}(r_{ia}, r_{ib}),
\]

\[
H_{\text{HY-HY}} = \sum_{a=1}^{N_{\text{HY}}} \left( \sum_{i=1}^{s_a} \frac{p_{ia}^2}{2m_{ia}} \right) + \left[ \sum_{a=1}^{N_{\text{HY}}} \sum_{b=1, b \neq a}^{N_{\text{HY}}} \left( w(X_a)w(X_b) V_{\text{AT}}(r_{ia}, r_{ib}) \right) \right] + \left[ \sum_{a=1}^{N_{\text{HY}}} \left( 1 - w(X_a)w(X_b) V_{\text{CG}}(R_{ia}, R_{ib}) \right) \right],
\]

\[
H_{\text{AT-HY}} = \sum_{a=1}^{N_{\text{AT}}} \left( \sum_{b=1, b \neq a}^{N_{\text{HY}}} V_{\text{AT}}(r_{ia}, r_{ib}) \right),
\]

\[
H_{\text{CG-HY}} = \sum_{a=1}^{N_{\text{CG}}} \left( \sum_{b=1, b \neq a}^{N_{\text{HY}}} V_{\text{CG}}(R_{ia}, R_{ib}) \right),
\]

\[
H_{\text{intra}} = H_{\text{bond}} + H_{\text{angle}},
\]

where \(H_{\text{bond}}\) is the sum of all intramolecular atom-atom bonding energies, and \(H_{\text{angle}}\) is the sum of all the intramolecular bonding angle energies. The specific form of \(H_{\text{intra}}\) depends on the molecular model used and it is defined in Sec. IV.

In the formulas above we have assumed that the atomistic degrees of freedom for the kinetic energy and of the intramolecular potentials are present in all molecules independently of the resolution (double resolution everywhere) but their adaptive character is considered only in relation to the intermolecular interaction sites. Such a situation corresponds to the actual (current) numerical implementation of the adaptive resolution scheme; in fact from the numerical point of view the calculations of the forces correspond to about
75% of the computational effort. Moreover, the coarse-grained potential is derived to reproduce the thermodynamics and the probability distribution function in space up to the two-body case, i.e., the radial distribution function $g(r)$, of the atomistic resolution. At this point we can rewrite Eq. (5) as the sum of “exact” and “perturbation” terms as follows. The exact terms involve all contributions to the energies and interactions in Eq. (12) that are well defined by the physics and do not involve any external or artificial quantities, such as $w(x)$. Thus,

$$H_{\text{exact}} = [H_{\text{AT}} + H_{\text{AT-HY}}] + [H_{\text{CG-CG}} + H_{\text{HY}}],$$

where $K_{\text{HY}} = \sum_{i=1}^{N_{\text{HY}}} \left( \sum_{j=1}^{N_{\text{HY}}} \frac{\mathbf{p}_i \cdot \mathbf{p}_j}{2m} \right)$ is the kinetic energy of the hybrid region. The remaining perturbations to the exact Hamiltonian related to the presence of the hybrid region read

$$\Delta H = \sum_{i=1}^{N_{\text{HY}}} \left( \sum_{j=1}^{N_{\text{HY}}} w(x_a)w(x_b)V_{\text{AT}}(\mathbf{r}_a, \mathbf{r}_b) \right)$$

$$+ \sum_{i=1}^{N_{\text{HY}}} \left( \sum_{j=1}^{N_{\text{HY}}} (1 - w(x_a))w(x_b)V_{\text{CG}}(\mathbf{R}_a, \mathbf{R}_b) \right).$$

Let us now reformulate the definition of the switching function $w(x)$ in terms of some characteristic lengths of the problem:

$$w(x) = \tilde{w}(\phi(x)/d) = \tilde{w}(\xi),$$

where $\phi(x)$ is the signed distance from the boundary of the atomistic region and $d$ is the characteristic thickness of the hybrid or transition region; i.e., the hybrid region is covered, letting $0 \leq \xi < 1$. One of the characteristic lengths of primary importance is $l_c$, that is, the range of molecular interaction. Other characteristic lengths are the size of the AT region, $D$, and that of the CG region, $L$.

To proceed in the development of an asymptotic limit formulation, and to meet the actual numerical setup used in AdResS, we assume to work in a regime with $\epsilon = l_c/d \ll 1$. We observe that the atomistic and coarse-grained interatomic forces, $F_a^{\text{AT}}$ and $F_a^{\text{CG}}$, are very small when $|X_a - X_b| > l_c$. As a consequence, the weighting functions $w(x_a), w(x_b)$ in the expression for forces in the hybrid region can only contribute sizably when $|X_a - X_b| \lesssim l_c$. Thus, whenever the intermolecular forces are sizable, a Taylor expansion of the weighting function evaluated at $\tilde{X} = \frac{X_{\text{AT}} + X_{\text{HY}}}{2}$ yields, e.g., for $X_a$,

$$w(x_a) = w(\tilde{X}) + \epsilon \frac{d \tilde{w}}{d \xi} X_a - \tilde{X} \nabla \phi(\tilde{X}) + O(\epsilon^2)$$

and the perturbation Hamiltonian reads

$$\Delta H = \Delta H(\epsilon) = w(\tilde{X})^2 V_{\text{AT}} + [1 - w(\tilde{X})^2] V_{\text{CG}} + O(\epsilon).$$

Since $w(\tilde{X})$ is slowly varying, i.e., $\frac{d w(\tilde{X})}{d x} = O(\epsilon)$, we obtain the leading force term,

$$F_{\text{HY}} = \tilde{w}(\tilde{X})^2 \tilde{F}_{\text{AT}} + [1 - \tilde{w}(\tilde{X})^2] \tilde{F}_{\text{CG}} + O(\epsilon).$$

This is a force formulation for the transition region, which is close to the formulation known from the standard adaptive resolution approaches discussed before; force formulation, until now, was not derived from an Hamiltonian point of view. Here, this force emerges as the gradient of the interpolated potential up to perturbations of order $\epsilon$. As a consequence, the particle motion within the transition region follows some perturbed weakly non-Hamiltonian dynamics. This is the starting point of a further asymptotic analysis regarding different scales involved in the problem; in particular, it becomes of interest to perform such an analysis with respect to the other characteristic lengths involved, such as $D$ and $L, D$, the size of the AT region, represents in principle the smaller scale; that is, it is the region involved in the observation of a very local event, looked at with all atomistic details. Instead $L$, the extension of the coarse-grained region, represents the scale of the part of the system with the role of a large macroscopic reservoir which ensures that the macroscopic quantities of the thermodynamics state (particle density, temperature, and pressure) are preserved. In the following we explain how to use Eq. (17) and the asymptotic approach for $L, d$, and $D$ to determine the degree of the perturbation and thus, if this is negligible, to identify the conditions under which the adaptive scheme is essentially Hamiltonian.

### B. Hamiltonian versus dissipative scheme

Let us define the quantities $\eta = \frac{D}{T}, \zeta = \frac{\epsilon}{T}$, and $\lambda = \frac{D}{T}$. We can perform numerical simulations using Eq. (17) and sample the space of $\eta$, $\zeta$, and $\lambda$ to check numerically combinations of such parameters for which the system is dissipative or Hamiltonian. The ideal setup, from the mathematical point of view, for a Hamiltonian-like behavior, would be $\eta \ll 1$, $\zeta \ll 1$, and $\lambda \ll 1$. In fact one could have a large coarse-grained region which, being dominant, ensures that the overall thermodynamic conditions are preserved, and a large HY region compared to the AT region where the condition $\frac{d w(\tilde{X})}{d \xi} \ll 1$ can be reasonably met. Moreover, such a situation is also optimal for the physical interpretation of multiscale analysis, which is a very localized event analyzed in a very small (AT) region compared to the rest of the system. However, since the asymptotic considerations described above rely predominantly on the smallness of $\epsilon = l_c/d$, i.e., on the hybrid region being “thick” in comparison with the molecular interaction distance, the case of a large $D$ and a small $d$ would be equally acceptable from this point of view as long as $\epsilon \ll 1$ is guaranteed. In this regime, the large coarse-grained and atomistic regions are expected to produce more reliable spatially nearly homogeneous statistics, although related simulations will, of course, be substantially more expensive. In the numerical simulation we have as a reference the full atomistic simulation in a microcanonical ensemble (NVE); that is, the system is self-contained and there is no external thermostat (as it is instead in the NVT). We compare the full atomistic simulation with the adaptive resolution simulation; the latter is also performed without the help of any external tool or thermostat. By monitoring the temperature as a function of time we check whether the system behaves strongly in a dissipative way or closer to conservative; if the system is dissipative then it is certainly
V\_b(r) = \frac{1}{2} k\_b (r - b\_0)^2
\quad \text{and}
V\_a(\theta) = \frac{1}{2} k\_a (\theta - \theta\_0)^2

FIG. 2. Intramolecular interactions of the full atomistic model.

not Hamiltonian. If the system is conservative we then check whether structural properties of the atomistic region are the same as those of a full atomistic simulation and structural

\[ f\_a = \begin{cases} \frac{1}{2} - C\_a, & r < r\_1 \\ \frac{1}{2} - s \left( r - r\_1 \right)^3 - \left( r - r\_1 \right)^4 - C\_s, & r\_1 \leq r < r\_c \\ 0, & r \geq r\_c \end{cases} \]

where \( s \), \( r\_1 \), and \( r\_c \) are chosen such that \( f\_a(r\_c) = 0 \), \( f\_a'(r\_c) = 0 \), and \( f\_a''(r) \) is continuous at \( r = r\_1 \). So the original LJ interaction is changed smoothly from \( r = r\_1 \) to zero at \( r = r\_0 \). \( C\_s, k\_a, k\_b, b\_0 \), and \( b\_0 \) are taken from the Gromacs OPLS/AA force field [24,25].

The choice of such a molecular model is justified by the fact that we need a model sufficiently simple for a straightforward numerical implementation but at the same time also sufficiently complex to represent a valid challenging test for the mathematical principles behind the method. With this model, we can now write the specific form of \( H\_\text{intra} \):

\[ H\_\text{intra} = \sum\_a \left( \frac{4}{3} k\_a \left( \| \textbf{r}\_a - \textbf{r}\_\text{H}_a \| - b\_0 \right)^2 + \sum\_j \sum\_i \frac{4}{3} k\_a \left( \theta\_a^{\text{H}_a\text{C}_a\text{H}_j} - \theta\_0 \right)^2 \right) \]

this choice is taken for practical purposes so that our results can be directly applied to any atomistic simulation of typical molecular liquids of interest in chemical physics. In the next section we report the technical details of the simulations so that our results can be reproduced.

IV. NUMERICAL SIMULATION

A. Molecular model

We have constructed a toy model: a molecule with methanlike structure with tetrahedral arrangement of lighter atoms (hydrogenlike, thus named H) connected via flexible (harmonic) bonds to a central heavier atom (carbonlike, thus named C) and with the H-C-H bond angle that is also described by a harmonic potential (see Fig. 2). Intermolecular interactions are described by a smoothed Lennard-Jones (LJ) 12-6 potential [23]:

\[ V\_XY(r) = C\_1\_2 \_12 f\_12(r) - C\_6 \_12 f\_6(r), \]

where \( X \) and \( Y \) can be C or H (see Fig. 3).

\[ f\_12(r) = \begin{cases} \frac{1}{r^2} - C\_s, & r < r\_1 \\ \frac{1}{r^2} - s \left( r - r\_1 \right)^3 - s \left( r - r\_1 \right)^4 - C\_s, & r\_1 \leq r < r\_c \\ 0, & r \geq r\_c \end{cases} \]

\[ f\_6(r) = \begin{cases} \frac{1}{r^6} - C\_s, & r < r\_1 \\ \frac{1}{r^6} - s \left( r - r\_1 \right)^3 - \left( r - r\_1 \right)^4 - C\_s, & r\_1 \leq r < r\_c \\ 0, & r \geq r\_c \end{cases} \]

where \( s \) can be 12 or 6, \( r\_1 = 1.1 \) nm is the switching radius for the smoothing, and \( r\_c = 1.3 \) nm is the cutoff radius of interaction. \( A\_s, B\_s, \) and \( C\_s \) are chosen such that \( f\_a(r\_c) = 0 \), \( f\_a''(r\_c) = 0 \), and \( f\_a''(r) \) is continuous at \( r = r\_1 \). The original LJ interaction is changed smoothly from \( r = r\_1 \) to zero at \( r = r\_0 \). \( C\_s, k\_a, k\_b, b\_0 \), and \( b\_0 \) are taken from the Gromacs OPLS/AA force field [24,25].

The choice of such a molecular model is justified by the fact that we need a model sufficiently simple for a straightforward numerical implementation but at the same time also sufficiently complex to represent a valid challenging test for the mathematical principles behind the method. With this model, we can now write the specific form of \( H\_\text{intra} \):

\[ H\_\text{intra} = \sum\_a \left( \frac{4}{3} k\_a \left( \| \textbf{r}\_a - \textbf{r}\_\text{H}_a \| - b\_0 \right)^2 + \sum\_j \sum\_i \frac{4}{3} k\_a \left( \theta\_a^{\text{H}_a\text{C}_a\text{H}_j} - \theta\_0 \right)^2 \right) \]

where \( \textbf{r}\_a \) is the spatial coordinate of the carbon atom of the \( a \)th molecule, \( \textbf{r}\_a^{\text{H}_a} \) is the spatial coordinate of the \( i \)th hydrogen atom of the \( a \)th molecule, and \( \theta\_a^{\text{H}_a\text{C}_a\text{H}_j} \) is the bond angle \( \text{H}_a\text{C}_a\text{H}_j \) of the \( a \)th molecule.

We then perform numerical simulations at a temperature of 111 K and pressure 1.0 atm, i.e., a condition at which the structure of this system is typical of a standard liquid. The coarse-grained model is built by performing the iterative Boltzmann inversion (IBI) procedure with pressure correction [26]. This procedure ensures that the resulting coarse-grained model matches the c.m.-c.m. radial distribution function (RDF) of the full atomistic model and ensures the same pressure (at the same given temperature). A pictorial representation of the molecular coarse-grained model and the resulting numerical or tabulated potential are shown in Fig. 4.

At this point we possess all the necessary ingredients for performing adaptive resolution simulations using Eq. (17) in the \( d, D, L \) space. We keep \( l\_a \), that is, the range of molecular interaction, fixed at a value typical of atomistic simulations; properties of the coarse-grained region are the same as those of a full coarse-grained simulation; if the result is positive then, from a physical point of view, we can claim to have found a Hamiltonian that by allowing a spatial adaptive molecular resolution can preserve basic thermodynamic properties (i.e., temperature conservation) and structural properties of a full atomistic and full coarse-grained simulation.

B. Technical details

All simulations are performed using double-precision home-modified Gromacs 4.6.5 [27]. The IBI technique with pressure correction is employed to derive the coarse-grained model by using the VOTCA 2.6 package [28]. The time step for the IBI is 0.002 ps and the pressure correction is added every two in three iteration steps with a standard pressure smoothing at the post update stage. The smoothing potential for LJ interaction and for the coarse-grained potential is numerically implemented by using the user defined table with spacing 0.0005 nm. The switching radius of the smoothing is 1.1 nm while the cutoff radius \( l\_c \) is 1.3 nm.
FIG. 3. Intermolecular interactions of the full atomistic model.

Preparation and equilibration of the system

The system is prepared starting from an initial box containing 470 “methane” molecules with dimensions of the system $3.0 \, \times \, 3.0 \, \times \, 3.0 \, \text{nm}^3$. Next the box is copied along the $x$ direction to produce the initial configuration for larger systems. It follows a full atomistic $NPT$ simulation (that is, at fixed $0.5 \, \times \, 1.5 \, \times \, 2.5 \, \text{nm}^3$).

number of molecules, pressure, and temperature, while the volume is allowed to fluctuate) using the Parrinello-Rahman coupling method [29,30] for 50 ps in order to determine the equilibrium volume. The coupling type is isotropic, the reference pressure is 1.0 atm, the time constant for pressure coupling is $\tau_p = 2.0 \, \text{ps}$, and the compressibility is $4.1 \times 10^{-3} \, \text{bar}^{-1}$. Next, a $NVT$ simulation is thermally equilibrated, with $V$ fixed at the equilibrium value of the $NPT$ simulation, by using a Langevin thermostat for 100 ns. The reference temperature is 110 K and the time constant for temperature coupling is $\tau_t = 0.1 \, \text{ps}$. The $NVT$ thermally equilibrated configuration is used for a full $NVE$ atomistic simulation (reference calculation) and for an equivalent adaptive resolution simulation. The time step used for all the simulations is 0.0005 ps.

C. Results and discussion

Table I reports results about some representative examples of systems that we have studied by sampling the $D,d,L$ ($\eta,\zeta,\lambda$) space. It must be reported that $\epsilon$ is not the sole parameter governing the H (Hamiltonian) or NH (non-Hamiltonian) behavior. In fact the size of the coarse-grained and of the atomistic regions also play an important role. For the atomistic region

![Table I](image)

<table>
<thead>
<tr>
<th>Label</th>
<th>$N_{\text{tot}}$</th>
<th>$D$ (nm)</th>
<th>$d$ (nm)</th>
<th>$L$ (nm)</th>
<th>Behavior</th>
<th>$\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>47000</td>
<td>61.5</td>
<td>15</td>
<td>73.1</td>
<td>H</td>
<td>0.087</td>
</tr>
<tr>
<td>2</td>
<td>47000</td>
<td>13.5</td>
<td>3</td>
<td>133.5</td>
<td>NH</td>
<td>0.433</td>
</tr>
<tr>
<td>3</td>
<td>28200</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>H</td>
<td>0.043</td>
</tr>
<tr>
<td>4</td>
<td>20680</td>
<td>5</td>
<td>30</td>
<td>31</td>
<td>H</td>
<td>0.043</td>
</tr>
<tr>
<td>5</td>
<td>18800</td>
<td>15</td>
<td>15</td>
<td>30</td>
<td>H</td>
<td>0.087</td>
</tr>
<tr>
<td>6</td>
<td>15980</td>
<td>5</td>
<td>15</td>
<td>31</td>
<td>H</td>
<td>0.087</td>
</tr>
<tr>
<td>7</td>
<td>12220</td>
<td>5</td>
<td>5</td>
<td>29</td>
<td>NH</td>
<td>0.260</td>
</tr>
<tr>
<td>8</td>
<td>12220</td>
<td>5</td>
<td>15</td>
<td>19</td>
<td>NH</td>
<td>0.087</td>
</tr>
</tbody>
</table>

![FIG. 4](image)

FIG. 4. (a) Coarse-grained model and its corresponding numerical potential derived by the IBI procedure. (b) Comparison between the c.m.-c.m. radial distribution function of the full atomistic simulation and the equivalent obtained from the coarse-grained potential of (a) within the IBI procedure.

![FIG. 5](image)

FIG. 5. Temperature as a function of time for system 6. Data are analyzed after 500 ps to allow the system time for basic equilibration in the adaptive setup.
one would like a size as small as possible (yet statistically valid) so that the saving of the computational resources can be optimized. For the coarse-grained region instead a small size would mean that the distribution of the dissipation of the hybrid region will be distributed among a relatively small number of molecules (that is, the coarse-grained region does not act as a large “reservoir”). As a consequence if the coarse-grained region is too small, the perturbation of the hybrid region will induce a large perturbation (per molecule) of the coarse-grained region and thus a sizable perturbation to the overall thermodynamics of the system. Systems of smaller size (i.e., with fewer than 12,000 molecules) are not reported because they do not meet either the mathematical or the (intuitive or practical) physical conditions for being Hamiltonian. Instead the systems reported in Table I are in principle all possible candidates for being Hamiltonian systems; however, some meet the conditions better than others. The numerical indicator for the classification in H or NH as reported in Table I is that in the time frame between 0.5 and 1.0 ns the temperature of the adaptive systems overlaps with that of the full atomistic simulation of reference, and that structural properties in the atomistic region (and trivially in the coarse-grained region) agree with those of the corresponding full atomistic simulation of reference. This classification is based on the idea that since after 0.5 ns the system is equilibrated, for (at least) short trajectories of 0.5 ns (i.e., from 0.5 to 1.0 ns) we run in a Hamiltonian regime which fully corresponds to that of full atomistic simulation. Thus for calculating statistical properties one may create an ensemble of independent trajectories of (at least) length 0.5 ns and obtain valid results by averaging over the ensemble. In previous adaptive resolution studies such an approach (even over shorter time windows) has been already used [31]. However, in Sec. IV C we also discuss one example where the simulation run is much longer and analyze how the dissipation due to the perturbation in the hybrid region influences the actual numerical results.

In general, some systems, in the time window considered, have shown a clear Hamiltonian nature, and they are characterized by a relatively large HY region and thus the condition $\frac{d\varepsilon_\mathrm{HY}}{dt} \ll 1$ is satisfactorily met and the CG region is large enough to provide thermodynamic stability to the rest of the system. Figure 5 shows the temperature as a function of time compared with the equivalent full atomistic NVE simulation for system 6, taken as a representative example; the agreement is remarkable (if the system was dissipative the temperature...
would show a sizable drift compared to the reference full atomistic simulation). As discussed before, Fig. 5 is not sufficient to justify the claim that the Hamiltonian of the system is a “proper” adaptive Hamiltonian. In fact, it may well be that while the system preserves the temperature in a proper way, other properties are instead modified in such a way that they do not reproduce the properties of the reference full atomistic system (or full coarse-grained system). For this reason we studied two important structural properties, that is, the molecular number density and the radial distribution functions, C-C, C-H, and H-H in the atomistic region; here we report the result for the most critical case that is for system 6. Figure 6 shows for system 6 the density across the box after 750 ps. After such time if the system did not behave properly it should display large deviations from the reference NVE results. The density profile agrees with the reference one, with differences, in the most unfavorable case, of about 6% in the region of maximal perturbation, that is, in the HY region; however, in the rest of the system the differences are below 2% which is highly satisfactory. Moreover, Figure 7 shows that the atomistic radial distribution functions of the adaptive systems essentially overlap with those of the reference full atomistic system. Other systems reported in Table I and characterized as “H” show the same accuracy of system 6.

On the basis of this result we can claim that we have found several systems where the adaptive setup corresponds closely to a self-contained Hamiltonian which produces highly satisfactory results on time windows of at least 1000 ps. As said before, the time window considered is the same as that used in other adaptive work to calculate static properties of the full atomistic region. Instead systems 2 and 7 do not produce satisfactory results; in system 2 the transition region is very small and thus it strongly violates the mathematical condition of $w(x)$ being slowly varying. In fact despite the fact that the dissipation produced by the sharp transition can be mostly adsorbed by the large AT and very large CG regions, the dissipation of energy is clearly visible in the plot of the temperature as a function of time (see Fig. 8). Finally for system 7 we have smaller AT and CG regions than for system 2; this, together with the small size of the HY region, makes the system highly dissipative as reported in Fig. 9. System 8 instead, although it has a relatively large HY region, is characterized by a CG region which is too small to act as an equilibrating reservoir.

D. Hamiltonian behavior versus dissipative behavior: Beyond 1.0 ns

In the previous section we have chosen a rather strict criterion for defining a system’s behavior as H or NH. However, the strict criterion allows us to state that in the time frame considered there are no differences between a full atomistic simulation and an adaptive resolution simulation. Nevertheless, one needs to address the question of what happens beyond the time window considered, above all for systems classified as H; in fact, the loss or gain of energy is cumulative, that is, it adds up during the simulation. This process is inevitable since, besides the integration error present...
also in the full atomistic simulation, in the adaptive resolution simulation we have only an approximate Hamiltonian system. In this section we discuss this aspect by considering the three systems discussed in the previous sections. In particular for system 6, we have performed a much longer simulation (4.0 ns) and compared its results with the equivalent full atomistic simulation. The slopes of the linear regression of the temperature versus time for systems 2 and 7 are respectively \(6.6 \times 10^{-4}\) and \(1.15 \times 10^{-3}\), that is, the kelvin dissipated (acquired) per picosecond. This implies that in a time window of 4000 ps (4.0 ns), system 2 increases the temperature by about 2.7 K while system 7 increases the temperature by 4.6 K. The deviations found are not dramatic, but certainly sizable. For system 6 instead, being classified as H we have run a longer simulation and the curve of the temperature versus time is reported in Fig. 10.

The slope in this case is \(4.4 \times 10^{-4}\) and this should be also compared with the correspondent quantity of the full atomistic simulation which is \(1.0 \times 10^{-3}\). The result is that the deviation from the ideal (target) temperature after 4.0 ns is about 1.6 K while the difference with respect to a corresponding full atomistic simulation is about 1.3 K. The question is now whether for system 6 one may in practice use the longer simulation for calculating, e.g., structural properties.

In Figs. 11 and 12 we have calculated the density across the system and various atomistic radial distribution functions in the atomistic region for a time frame between 1.0 and 4.0 ns. This time frame considers times when the curve of the temperature of the adaptive resolution system and of the full atomistic system are not strictly overlapping anymore. The results show that the agreement of the structural properties is still satisfactory. For longer time windows of course system 6 will start to deviate significantly from a Hamiltonian behavior. In any case, the results of this section show the robustness of the approach for time windows which can be used for productive runs; certainly this approach represents a promising basis for building an improved numerical algorithm.
V. DISCUSSION AND CONCLUSION

Hamiltonian-based approaches which introduce the idea of adaptive molecular resolution are gaining popularity in the community of molecular dynamics. The conceptual background on which they are based is still a subject of dispute; in particular the existence of a global adaptive resolution Hamiltonian written solely in terms of particles’ degrees of freedom is still an open question [1,14,22,32]. In this work we have proposed a modification of the coupling between the atomistic and the coarse-grained region; the key difference from the other methods [4,5] is that the Hamiltonian term which couples the atomistic region with the hybrid region and the Hamiltonian term which couples the coarse-grained region with the hybrid region are written in the form of full atomistic interactions and full coarse-grained interactions, respectively, without the introduction of any space-dependent weight. This choice allows us to write the Hamiltonian corresponding to the interactions between the molecules of the hybrid region as a perturbation (introduced by a space-dependent interpolating weighting function) to the sum of a full atomistic and a full coarse-grained Hamiltonian. The perturbation then is a well-defined term, independent of the interactions with the other regions, and thus it can be mathematically analyzed in an asymptotic approach. From the technical point of view the partitioning proposed offers the possibility to implement the numerical procedure in a simplified and more general way for any molecular dynamics code: the atomistic and coarse-grained regions can be treated independently and then synchronized by adding the perturbation term of the Hamiltonian. We have carried out an asymptotic analysis and a numerical verification of the proposed scheme and show that when some mathematical conditions are reasonably fulfilled the system behaves practically as a Hamiltonian system. It should be underlined that the mathematical conditions are such that for molecular simulations in the current implementation the setup required by the Hamiltonian approach implies the choice of large systems and thus the necessity to perform large expensive calculations when compared to other adaptive Hamiltonian-like approaches [4,5] or to (technically) equivalent non-Hamiltonian or stochastic adaptive schemes based on the idea of an open boundary or grand-canonical-like approach [6–9,17–19,33]. We hope that the conceptual and (potentially) technical advantages offered by our partitioning together with the detailed mathematical and numerical analysis carried out here can be employed for improving the numerical and conceptual development of the adaptive resolution technique.

ACKNOWLEDGMENTS

This research has been funded by Deutsche Forschungsgemeinschaft (DFG) through Grant No. CRC 1114 (project C01) for R.K. and L.D.S. and by the China Scholarship Council (Grant No. 201506010012) for J.L.Z.