Physics Reports 693 (2017) 1–56

Contents lists available at ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

Molecular systems with open boundaries: Theory and simulation

Luigi Delle Site [a,*], Matej Praprotnik [b,c,**]

[a] Institute for Mathematics, Freie Universität, Arnimallee 6, D-14195 Berlin, Germany
[b] Department of Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia
[c] Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:
Accepted 9 May 2017
Available online 22 June 2017
Editor: H. Orland

Keywords:
Open molecular systems
Open boundary simulations
Grand Canonical ensemble

ABSTRACT

Typical experimental setups for molecular systems must deal with a certain coupling to the external environment, that is, the system is open and exchanges mass, momentum, and energy with its surroundings. Instead, standard molecular simulations are mostly performed using periodic boundary conditions with a constant number of molecules. In this review, we summarize major development of simulation methodologies, which, contrary to standard techniques, open up the boundaries of a molecular system and allow for exchange of energy and matter with the environment, in and out of equilibrium. In particular, we construct the review around the open boundary simulation approaches based on the Adaptive Resolution Scheme (AdResS), which seamlessly couples different levels of resolution in molecular simulations. Ideas and theoretical concepts used in its development lie at the crossroad of different fields and disciplines and open many different directions for future developments in molecular simulation. We examine progress related to theoretical as well as novel modeling approaches bridging length scales from quantum to the continuum description and report on their application in various molecular systems. The outlook of the review is dedicated to the perspective of how to further incorporate rigorous theoretical approaches such as the Bergmann–Lebowitz and Emch–Sewell models into the molecular simulation algorithms and stimulate further development of open boundary simulation methods and their application.

Contents

* Corresponding author.
** Corresponding author at: Department of Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia.
E-mail addresses: luigi.dellesite@fu-berlin.de (L. Delle Site), praprot@cmm.ki.si (M. Praprotnik).

http://dx.doi.org/10.1016/j.physrep.2017.05.007

1. Introduction

In literature, the term *"open systems"* refers to systems exchanging energy and matter with an external environment. However, in most of the cases, state-of-the-art and traditional literature considers only the exchange of energy between a system and its environment [1–7]. Indeed, externally driven exchange of energy via an external device is common both in experiments and construction of technological tools, but it is certainly neither the most common in nature nor the most advanced procedure in technology. Besides, it is unlikely that naturally-occurring systems have closed boundaries and that the exchange with an external environment can be reduced to a thermalization process/exchange of energy, only. Hence, the development of modern technologies is not to be restricted to systems capable of exchanging only energy, but actually spectacular progress has been made by allowing systems to also exchange matter with the outside. There are many examples, spanning different fields of physics, different scales, different experimental and computational techniques. Chemical reactions in a biochemical environment do not necessarily conserve the number of molecules [8,9], ion fluxes through nanopores and transmembrane channels are natural cases of exchange of matter between different compartments [10] and a key technology process is the adsorption of guest molecules in microporus materials [11]. Furthermore, liquid–vapor condensation can be properly described as two different systems exchanging molecules [12], the addition of colloids in a solvent is the most efficient experimental technique to study the phase diagram of polydisperse colloidal materials [13], nucleation is an activated process with exchange (capture) of matter between the nucleation site and the environment [14] and this is directly linked to the emerging concept of nanothermodynamics and its applications [15–17]. Also, squeezing and shearing of nanoconfined fluids for biolubrication, nanotribology, and surface engineering are properly described in terms of open systems which exchange particles with a reservoir at fixed chemical potential [18].

The few illustrative examples, mentioned above, concern condensed matter cases at classical scales. However, the process of exchange of matter is becoming also of major interest in quantum systems [19–22] and corresponding technologies. Traditional research in quantum open systems with exchange of matter has interested the field of microelectronics for a long time. The scattering of electrons in a resistor by phonons and the escape of the electron from the resistive material to the conductive contact is a standard process in building electrical microcircuits [23] and MOSFET devices [24]. Furthermore, electronic structure of open systems plays a key role in the computation of redox potentials in concentration cells, where the electrodes can be modeled as electron reservoir [25], and in rational compound design, with the sampling of the chemical space in a molecular Grand Canonical (GC) ensemble [26,27]. Moving forward, a field in great expansion and of potentially groundbreaking technological implications is that of cold gases in optical traps. Condensate fluctuations of a Bose gas confined in an optical trap can be viewed as a subsystem which exchanges particles with a larger environment [28–30]. Interestingly, for an optically confined quantum gas the effect of particle loss due to, for example, a collision process, where

the released energy is much larger than the depth of the trap, is described via master equations regulating the change of number of particles [31]. This led to the formulation of the so called continuous quantum Zeno effect. In general, the quantum Zeno effect refers to the fact that repeated measurements on a quantum system inhibit coherent processes [32]. The extension mentioned above (continuum) refers to the fact that a system in contact with an environment is subjected to a continuous measurement from the environment, changing the number of particles [33]. The problem of inhibiting coherence, in turn, is of great importance in the extremely fashionable field of quantum computers with harboring hopes for a new scientific revolution [34].

The overview, given above, leads to the conclusion that progress in the development of rigorous theoretical frameworks and corresponding accurate and efficient computational tools are a timely necessity. In this perspective, the intention of this review is to provide an overview of the state of the art in theory and simulation of open boundary systems. The core of the presentation will be the treatment of molecular systems with open boundaries in molecular simulation. This choice is dictated by the rather broad field of applications of molecular simulation and the large number of practitioners in the field. In particular, the review is constructed around the Adaptive Resolution Simulation method (AdResS) since it bridges theoretical models and computational techniques, where contributions from different fields converge and, possibly, branch out. As a result, particular attention will be dedicated to the actual and potential connection between rigorous theoretical principles and technical/computational implementations.

The paper is organized as follows: Section 2 is dedicated to the pure theoretical treatment of systems with open boundaries. The basic ideas regarding the required mathematical formalization are illustrated through some specifically treated examples. Section 3 is dedicated to the molecular simulation of open systems. First, we provide a general and qualitative overview of classical molecular simulation methods for simulations in the GC ensemble. Subsequently, we pass to the description of the coupling technique in the adaptive resolution fashion and analyze the various branching directions that this approach is taking. Several features of the technique can be already directly linked to the formalization of Section 2 and a corresponding discussion of such connections is presented. In Section 4, taking the classical case as a reference, we address the natural question of a quantum mechanical treatment of molecular systems with open boundaries. For this purpose, we provide a basic overview of methods and techniques available to treat the problem at quantum level. Because of the higher complexity of quantum mechanics compared to classical mechanics, the state-of-art in the quantum treatment is not yet at the level of the classical case. Nevertheless, we underline that some direct extensions of classical techniques/ideas, e.g. AdResS, to the quantum case can be already done. In Section 5, we address the natural question of a possible extension of open boundary techniques to non-equilibrium situations. For illustration, we discuss two recent examples of satisfactory application. Finally, in Section 6, we discuss the potential connections between the various mathematical/theoretical models and computational techniques, we have reported in this review, in the perspective of future developments.

2. Theoretical approaches and their mathematical formalization

An interesting point about pure theoretical/mathematical treatment of the problem of open boundary systems is that most of the papers, in which this problem is treated, deal with non-equilibrium statistical mechanics. Thus, the possibility of exchange of particles between a system and a reservoir is anyway seen as a sub-case of a system in non-equilibrium, as for example a unidirectional current of particles. This point of view is understandable because typical problems of interest are those where current fluxes are produced by a source and then injected into a system, as for example in electric devices, see e.g. [23]. However, in this section we will see that the condition of flux balance (necessary for the equilibrium), although trivial to formalize, is not trivial when it comes to the explicit specification of the action of operators or functions that allows for such a balance. In fact, this is, in our view, the major challenge in modern research, above all for quantum systems, as will be discussed in the final sections of this review.

In general, there are two main approaches to describe the interaction between a system and its reservoir: (a) impulsive interaction, i.e., an instantaneous interaction, which produces a (stochastic) discrete transition in number of particles for the system; (b) coarse-grained interaction, where the reservoir enters into the dynamics of the system only through macroscopic quantities. The latter are, from a statistical point of view, time-independent but derived from the underlying microscopic degrees of freedom (DOFs) of the reservoir. In the following, we present two examples of methods which are in categories (a) and (b), respectively. The first one is the approach due to Bergmann and Lebowitz [35] and the second one is due to Emch and Sewell [36]. In our view, these two approaches are prototypes that provide very general and clear examples of the problems (and possible solutions) arising when treating systems, which exchange particles.

2.1. Bergmann–Lebowitz as the Grand Ensemble theoretical framework with stochastic reservoir

Bergmann and Lebowitz (BL) [35,37] (and later on Lebowitz and Shimony [38]) have discussed the statistical mechanics of open boundary systems and, in particular, have proposed the extension of Liouville's equation for systems that can exchange matter with a reservoir. The essential feature of the model is the definition of the action of the reservoir. The physical principle, which regulates such action, states that the interaction between the system and the reservoir is impulsive (stochastic). It is characterized by a discrete transition of the system from a state with M particles (X_M) to one with N particles (X'_N). The key point is that the macroscopic state of the reservoir is not influenced by the system. Thus, its microscopic DOFs do not need to be considered and particles entering into the system from the reservoir have, in thermal equilibrium, velocities

consistent with the temperature of the reservoir. An explicit back-reaction from the system to the reservoir, that changes the macroscopic (thermodynamic) state of the reservoir, is therefore not considered.

The model is general and considers the case of a system in contact with several independent reservoirs. The total action corresponds to the sum of the individual action of each reservoir. For simplicity, we will report the case of a single reservoir, because such a case will be discussed also later on in numerical implementations. From the formal point of view, a transition from a state X_M to a state X'_N is governed by a contingent probability $K_{NM}(X'_N, X_M)dX'dt$ where the kernel $K_{NM}(X'_N, X_M)$ is a stochastic function independent of time. $K_{NM}(X'_N, X_M)$ corresponds to the probability per unit time that the system at X_M has a transition to X'_N due to the interaction system-reservoir. The total interaction between the system and the reservoir corresponds to: $\sum_{N=0}^{\infty}\int dX'_N[K_{MN}(X_M, X'_N)\rho(X'_N, N, t) - K_{NM}(X'_N, X_M)\rho(X_M, M, t)]$. It follows that the general equation of time evolution of the probability, $\rho(X_M, M, t)$, is:

$$\frac{\partial\rho(X_M, M, t)}{\partial t} = \{\rho(X_M, M, t), H(X_M)\} + \sum_{N=0}^{\infty}\int dX'_N[K_{MN}(X_M, X'_N)\rho(X'_N, N, t) - K_{NM}(X'_N, X_M)\rho(X_M, M, t)]. \quad (1)$$

Then, under the condition of flux balance:

$$\sum_{N=0}^{\infty}\int dX'_N[K_{MN}(X_M, X'_N)\rho(X'_N, N, t) - K_{NM}(X'_N, X_M)\rho(X_M, M, t)] = 0. \quad (2)$$

It follows that the stationary solution for $\rho(X_M, M)$ is the density of the GC ensemble: $\rho_M(X_M, M) = \frac{1}{Q}e^{-\beta H_M(X_M)+\beta\mu M}$ where $\beta = 1/k_BT$ and μ the chemical potential. It is important to notice that Bergmann and Lebowitz have shown that the condition of flux balance is necessary and sufficient for having the GC distribution as the stationary solution. A different hypothesis regarding the action of the reservoir, based on the coarse-graining of its underlying microscopic details, has been instead considered in other work, as reported below.

2.2. Dynamic equation of system plus reservoir via the projector technique in Liouville space

The hypothesis of impulsive interaction between the system and the reservoir (“stosszahlansatz”), which is at the basis of the BL model, did not require the explicit knowledge of the microscopic evolution of the reservoir. In such a case, the treatment of a subsystem of a larger system implies the loss of microscopic information about the rest of the system. Instead, Emch and Sewell (ES) considered the case of a subsystem (S) in a larger system (reservoir/R) [36]. Their aim was to derive microscopic equations of motion for S from the microscopic equations for S+R. The relevant information for the dynamics of S is provided by properly filtering the action of the microscopic DOFs of R through projector operators (Zwanzig projector approach). The starting point is the quantum Liouville equation for the S+R system. Next, the equation is reduced to the equation of motion of the reduced statistical operator, which describes all the observables of S and the macroscopic variables of R (indicated with Σ_R). The derived equation is then reduced to the equation governing only S by making some hypothesis about the microscopic structure of R and its coupling to S. The final resulting equation is a self-contained dynamical equation for S, where the variables of R appear as averages quantities. The essential hypothesis behind such a reduction is that the reservoir action on S is done only via intensive properties. Here, it is assumed that although the extensive variables Σ_R could change by a finite amount due to the back-action of S, the intensive variables Σ_R/N become constants of motion for $N \to \infty$; $V \to \infty$ (thermodynamic limit). Thus, the evolution of S does not contain any time dependence of Σ_R. The fact that they explicitly treat the quantum mechanical case gives to their equations a general character since the classical equivalent can be derived from them. The qualitative description, given above, reports the essential physics involved in the model. Below, we will provide the basic points of the technical/mathematical derivation of the equations.

Let us define the Hilbert space of states of the total system S+R, $\mathcal{H}_{S+R}$, as the direct product of the Hilbert space of the system, $\mathcal{H}_S$, and of Hilbert space of the reservoir, $\mathcal{H}_R$: $\mathcal{H}_{S+R} = \mathcal{H}_S \otimes \mathcal{H}_R$. The description of the total system corresponds to the action of operators acting on $\mathcal{H}$. The variables of interest are the complete set of observables of S and a set of macroscopic observables of R. The macroscopic observables of R are indicated by the set, Σ_R, which is formed by intercommuting self-adjoint operators, $\{A_R\}$, acting the subspace of $\mathcal{H}_R$ corresponding to the pertinent microscopic description R. Each set of simultaneous eigenvalues of such operators, $\{A(E)\}$, defines a subspace $\mathcal{H}_E$ of $\mathcal{H}_R$; we will denote as, D_E the projector from $\mathcal{H}_R$ to $\mathcal{H}_E$. It follows that each macroscopic observable can be written as: $A_R = \sum_E A(E)D_E$, with $D_ED_{E'} = D_E\delta_{EE'}$; $D_E^* = D_E$; $\sum_E D_E = I_R$ (where I_R is the identity operator on $\mathcal{H}_R$). Next, the authors introduce the conceptual key point of their approach, they refer to it as: *“projector technique in Liouville space”*. Liouville space, $\mathcal{L}$ (which, since long, has had a broad application in physics [39]), is defined in such a way that $\mathcal{H}$ is associated with the set of all Hilbert–Schmidt operators, $\mathcal{L}$, which map $\mathcal{H}$ on itself: $\mathcal{L} \equiv \{A \in B(\mathcal{H}) \mid Tr(A^*A) < \infty\} \equiv \mathcal{L}(\mathcal{H})$, and the scalar product is defined: $(A, B) \equiv Tr(A^*B) \;\forall\; A, B, \quad \in \mathcal{L}$. Having defined the Liouville space, then, equivalently to the Hilbert space, let us define the Liouville space of the total system, $\mathcal{L}_{R+S}$, as the direct product of the Liouville space of the system, $\mathcal{L}_S$, and of the Liouville space of the reservoir, $\mathcal{L}_R$: $\mathcal{L}_{S+R} = \mathcal{L}_S \otimes \mathcal{L}_R$. We define the coarse-graining operator, $\mathcal{P}_R$, acting on $\mathcal{L}_R$ as:

$$\mathcal{P}_RB_R = \sum_E \langle B_R\rangle_E D_E;\ \forall B_R \in \mathcal{L}_R, \quad (3)$$

where $\langle B_R\rangle_E = Tr_R\frac{(B_R D_E)}{Tr_R D_E}$. In particular, in the view of a statistical description of the system, the formalism above is of particular relevance to the density operator ρ. In quantum mechanics, if the system is in a mixed state and the pure states $|\psi_i\rangle$ can be found with probability w_i, then the corresponding density operator is: $\rho = \sum_i w_i|\psi_i\rangle\langle\psi_i|$. In general, the density operator is represented through the density matrix, that is, if we choose an orthonormal basis, $\{|\phi_m\rangle\}$, the elements of the matrix are: $\rho_{mn} = \sum_i w_i\langle u_m|\psi_i\rangle\langle\psi_i|u_n\rangle = \langle u_m|\rho|u_n\rangle$. The ensemble average of an observable A is then defined though ρ as: $\langle A\rangle = \sum_i w_i\langle\psi_i|A|\psi_i\rangle = \sum_{mn}\langle u_m|\rho|u_n\rangle\langle u_n|A|u_m\rangle = \sum_{mn}\rho_{mn}A_{nm} = Tr(\rho A)$. In the formalism of the projector operator, defined before for every density operator ρ_R in $\mathcal{L}_R$, one has: $\mathcal{P}_R\rho_R = \sum_E Tr_R\frac{(\rho_R D_E)}{Tr_R D_E}$. It follows that for any macroscopic observables A_R and any state ρ_R one has: $\langle A_R\rangle_{\rho_R} = \sum_E Tr_R(\rho_R D_E)A(E)$; in this sense $Tr_R(\rho_R D_E)$ is the probability that the macroscopic observable A_R takes the values $A(E)$ when the system R is in the state ρ_R. At this point, we go back to the total system, S+R, and define the projector operator acting on it according to the variable of interest, that is no restriction imposed on S and only the set Σ_R: $\mathcal{P} = \mathcal{P}_R \otimes \mathcal{J}_S$, where $\mathcal{J}_S$ is the identity operator on $\mathcal{L}_S$. It follows that for any B_R in $\mathcal{L}_R$ and any B_S in $\mathcal{L}_S$ one has: $\mathcal{P}(B_R\otimes B_S) = \sum_E\langle B_R\rangle_E B_S^E$, with $\langle B_R\rangle_E$ as defined before and $B_S^E = D_E\otimes B_S$, with $Tr_R(B_R\otimes B_S) = (Tr_R B_R)B_S$. Now, we have all the ingredients to define: $\mathcal{P}\rho = \sum_E\frac{D_E}{Tr_R D_E}\otimes Tr_R(D_E\otimes I_S)\rho$. We can now replace ρ with an equivalent quantity which extracts from ρ the information of relevance to our description, that is the microscopic description of S and a macroscopic description of R. The technique of Zwanzig [40], based on the Laplace transform, is then used to pass from the standard von Neumann equation: $\frac{d}{dt}\rho = -iL\rho$ (with L the Liouvillian operator) to the equivalent (master) equation for $\mathcal{P}\rho$:

$$\frac{d}{dt}\mathcal{P}\rho(t) + i\mathcal{P}L\mathcal{P}\rho(t) + \int_0^t dt'\mathcal{P}L(I-\mathcal{P})\mathcal{U}(t-t')(I-\mathcal{P})L\mathcal{P}\rho(t') = 0, \tag{4}$$

with I the identity operator on $\mathcal{H}$ and $\mathcal{U}(t) = \exp[-i(I-\mathcal{P})L(I-\mathcal{P})t]$. Eq. (4) is derived under the assumptions that R and S are initially independent of each other, $\rho(0) = \rho_R(0)\otimes\rho_S(0)$ and that the initial state of R is given by the measurement of the set of macroscopic variable which implies $\mathcal{P}\rho(0) = \rho(0)$. Next, the authors show that $\mathcal{P}\rho(t) = \rho_R(0)\otimes\rho_S(t)$. Thus, Eq. (4) can now be written as a self-contained equation for the open system S:

$$\rho_R(0)\otimes\left(\frac{d}{dt} + iL_{eff}^S\right)\rho_S(t) = -\int_0^t dt', K(t-t')\rho_R(0)\otimes\rho_S(t'). \tag{5}$$

Finally, taking the trace with respect to R on both sides one obtains the master equation for $\rho_S(t)$:

$$\left(\frac{d}{dt} + iL_{eff}^S\right)\rho_S(t) = -\int_0^t dt'\mathcal{K}^S(t-t')\rho_S(t'), \tag{6}$$

where $\mathcal{K}^S(t)\rho_S(t) = Tr_R\{\mathcal{K}(t)\rho_R(0)\otimes\rho_S(t)\}$. The definitions of L_{eff}^S and $\mathcal{K}(t)$ are linked to another crucial point of particular interest to us, that is the possibility of defining the exchange of particles between R and S. In fact, one can separate the Hamiltonian of the system in three parts: $H = H_R + H_S + H_I$, i.e., the Hamiltonian of R, that of S, and a Hamiltonian of interaction between R and S, respectively. Correspondingly, we can define the related Liouville's operators L_R, L_S, L_I. Interestingly, the authors define $H_I = \int_{\Omega_R} dx\int_{\Omega_S} dyV(x,y)J_R(x)\otimes J_S(y)$, where x, y are the configuration coordinates of R and S respectively, Ω_R, Ω_S the volumes occupied, and $J_R(x)$, $J_S(y)$ are operators acting on $\mathcal{H}_r$, $\mathcal{H}_S$, respectively. They represent intensive variables such as particle number, for example, which could be function of the creation and annihilation operators for particles in R and in S. Their action is "coordinated" by $V(x,y)$, which is a potential of (direct) interaction between R and S. Thus, Eq. (6) is a general quantum-mechanical equation for a system S, which can exchange also matter with its environment. As a consequence of the definition above, $L_{eff}^S = L_S + L_I^S$, with $L_I^S\rho = [V_S,\rho]$ and $V_S = \int_{\Omega_S} dy\langle V(y)\rangle_0 J_S(y)$. The kernel is defined as follows, $\mathcal{K}(t) = \mathcal{P}\mathcal{U}_S(t)L_I(t)(I-P)\mathcal{U}'(t)L_I\mathcal{P}$, with $\mathcal{U}_S(t) = \exp\{-iL_S t\}$, $L_I(t) = \exp[i(L_R+L_S)t]L_I\exp[-i(L_R+L_S)t]$, and $\mathcal{U}'(t) = \exp\{-\int_0^t dt'(I-\mathcal{P}L_I(t'))(I-\mathcal{P})\}$. The classical limit of Eq. (6) can be made by identifying ρ_S with the classical distribution function of S in the phase space and by replacing the commutator with Poisson brackets. It must be reported that analogous equations for $\rho(0)_R\otimes\rho_s(t)$ via a projector operator technique were obtained also, for example, by Seke [41]. In Ref. [41], two different projector techniques were used, namely the Robertson method and Zubarev method [42–45]. However, as most of the papers dealing with open systems, differently from the approach of ES, the approach is limited to thermally open systems and the exchange of particles is forbidden by the choice of $\rho_s(t)$ as canonical density, which then is at the core of the derivation.

In conclusion, in our view, the treatment of ES offers a basic mathematical backbone for a theoretically rigorous approach to open boundary systems without the limiting hypothesis of a generic impulsive interaction between S and R (as in the BL model). On the other hand, the kernel term is rather complex and difficult to handle in practical calculations (differently, for example, from BL generalized Liouville equation of the previous section). Despite this obstacle, the ES model represents a path that needs to be explored more deeply in current research, above all for quantum mechanical systems, as it will be discussed in the section dedicated to the outlook and perspectives. In any case, regardless of the models (i.e., the BL or ES), once the equations for the probability density of the system are written, other aspects must be properly addressed. For example, a derivation of an analog of Liouville's theorem for systems with varying number of particles N. Moreover, one must address the question whether modifications to definitions, valid at fixed N, are required when N is varying. An example, reported in the next subsection, is that of the equilibrium time correlation functions with varying N.

2.3. Additional conceptual complications that arise when the particle number is not constant

Liouville's theorem states that a dynamical system of N particles in equilibrium conserves its distribution $\rho(\mathbf{q}, \mathbf{p})$, in the phase space, $(\mathbf{q}, \mathbf{p})$, along the trajectory. Such theorem justifies the statistical equivalence of points in the phase space, or, in mathematical language, the theorem reads: *the Lebesgue measure is preserved under the dynamics*. When N is constant everything is well established and the formulation is reported in any textbook of dynamics or statistical mechanics. However, for systems (in equilibrium), which exchange particles with external sources, the question has not been explicitly considered in standard literature. Recently, one of us has analyzed the problem and reached the conclusion that Liouville's theorem in a system with varying N can certainly be formulated for each, $N = \bar{N}$, canonical hyperplane [46]. The arguments and the corresponding derivation are summarized below. Let us start by analyzing the concept of conservation of Lebesgue measure which expresses Liouville's theorem when N is fixed:

$$\rho(\mathbf{q}_0, \mathbf{p}_0, 0) d\mathbf{q}_0 d\mathbf{p}_0 = \rho(\mathbf{q}_\tau, \mathbf{p}_\tau, \tau) d\mathbf{q}_\tau d\mathbf{p}_\tau. \tag{7}$$

$\mathbf{q}_0 = \mathbf{q}(0)$ (that is $\mathbf{q}(t)$ at $t = 0$) and same for $\mathbf{p}_0$. Similarly, it applies to $\mathbf{q}(t)$ and $\mathbf{p}(t)$ with $t = \tau$. We are in a situation of statistical equilibrium and thus $\rho(\mathbf{q}_0, \mathbf{p}_0, 0) = \rho(\mathbf{q}_\tau, \mathbf{p}_\tau, \tau)$. It follows that the compact formulation is:

$$d\mathbf{q}_0 d\mathbf{p}_0 = d\mathbf{q}_\tau d\mathbf{p}_\tau; \quad \forall \tau. \tag{8}$$

The derivation of Eq. (8) (see e.g. [47]) is based on the relation between $\mathbf{q}_0\mathbf{p}_0$ and $\mathbf{q}_\tau, \mathbf{p}_\tau$ through a coordinate transformation regulated by a Jacobian:

$$J(\mathbf{q}_\tau, \mathbf{p}_\tau, \mathbf{q}_0, \mathbf{p}_0) = \det(Q). \tag{9}$$

Here, Q is a $6N \times 6N$ matrix whose elements are:

$$Q_{ij} = \frac{\partial x_\tau^i}{\partial x_0^j}, \tag{10}$$

with $x_0 = (\mathbf{q}_1(0).....\mathbf{q}_N(0), \mathbf{p}_1(0).....\mathbf{p}_N(0))$ and equivalently $x_\tau = (\mathbf{q}_1(\tau).....\mathbf{q}_N(\tau), \mathbf{p}_1(\tau).....\mathbf{p}_N(\tau))$. The indices i, j label the $6N$ coordinates of x_0 and x_τ (i.e. $x^i = x^1....x^{6N}$ and equivalently for x^j; $(x^1, x^2, x^3) = (q_1^x, q_1^y, q_1^z)$ and $(x^{3N+1}, x^{3N+2}, x^{3N+3}) = (p_1^x, p_1^y, p_1^z)$ for example).

The key point is that if N is variable $\det(Q)$ cannot be formally calculated. This problem occurs because during the evolution in time the sets x_0 and x_τ do not necessarily have the same dimension. The question now is reduced to whether there exists a generalized principle, which extends the concept of Eq. (8) to the case of varying N. In Ref. [46], the BL model and its assumption of flux balance were considered:

$$\frac{d\rho(X_M, M, t)}{dt} = 0, \tag{11}$$

corresponding to:

$$\frac{\partial \rho(X_M, M, t)}{\partial t} = -\{\rho(X_M, M, t), H(X_M)\}. \tag{12}$$

Eq. (12) is formally equivalent to the standard Liouville's equation with fixed number of particles M with the crucial exception that now $\rho(X_M, M, t)$ and $H(X_M)$ are defined instantaneously only (w.r.t. M). The conclusion, that can be drawn from Eq. (12), is that a generalized Liouville's theorem for a system, which exchanges particles with a reservoir, can be written in the following form: *For systems in contact with a reservoir of particles, under the condition of statistical flux balance, the Lebesgue measure is conserved for each individual M*. The underlying conjecture is that a global Lebesgue measure does not exist, but it can be defined only for single subsets of the phase space each at fixed M. Our argument is that M is a discrete variable and its variations imply a discontinuous change of the phase space dimensionality; in our view this aspect represents a major obstacle. However, the proposal of Ref. [46] may be considered only as the simplest solution, of a local definition, where "local" indicates that the definition of a Lebesgue measure can be done at a given M:

$$d^N\mathbf{q}_0 d^N\mathbf{p}_0 = d^M\mathbf{q}_\tau d^M\mathbf{p}_\tau; \quad \forall \tau \quad where \quad M = N. \tag{13}$$

We notice that Peters in an unpublished paper [48] has suggested the concept of canonical hyperplanes. This essentially means that the equality applies whenever, after a time τ along a trajectory, the system returns to the same number of molecules M (from which the observation has started). In practice, it implies a process of sorting out instantaneous configurations with the same number of molecules M and for each subset apply the standard Liouville theorem. Together with Liouville's theorem, the concept of Liouville's operator and its involvement in the calculation of physical quantities of primary interest need to be revised when the number of particles is varying. Below, we report the case of equilibrium time correlation functions.

2.4. Equilibrium time correlation functions in an open boundary system

In Ref. [46], the so-called Kossakowski–Lindblad equation for the time evolution of the density matrix $\rho(t)$ was considered [49,50]:

$$\dot{\rho(t)} = L(\rho) = -i[H, \rho] + \frac{1}{2}\sum_j([L_j\rho, L_j^+] + [L_j, \rho L_j^+]), \tag{14}$$

where H is the Hamiltonian, L_j, L_j^+ the operators expressing the interaction of the system with a reservoir; they are called Lindblad operators. $\sum_j([L_j\rho, L_j^+] + [L_j, \rho L_j^+])$ makes Eq. (14) a rate equation (discontinuous jumps in the state of the system caused by the action of the reservoir). In fact, $[L_j\rho, L_j^+]$ and $[L_j, \rho L_j^+]$ can be interpreted as transition rates between two events. If the condition of flux balance: $\sum_j([L_j\rho, L_j^+] + [L_j, \rho L_j^+]) = 0$ is satisfied, as in case of a thermal bath (i.e., heat reservoir, being the context, in which usually Eq. (14) is used in literature), then the stationary solution for $\rho(t)$ is the density matrix of the canonical ensemble. In principle, the equation of ES discussed before falls in the category of Eq. (14), and, with particle exchange under the condition of flux balance, the stationary solution would be the density matrix of the GC ensemble. Such ideas will be discussed in a later section, explicitly dedicated to quantum systems. Instead, the classical analog of Eq. (14) is obtained by considering ρ as the probability distribution defined as $\rho(X_N, N, t)$, where X_N is a point in the phase space and N the total number of particles, and then substituting the commutator $[*, *]$ with the Poisson bracket $\{*, *\}$. Hence, the classical equivalent of Eq. (14) corresponds to the standard Liouville equation with the addition of the corresponding classical term related to the Lindblad operators. In such a view, the dynamic equation of the BL model discussed before falls under the category of (classical analog of) Eq. (14):

$$\frac{\partial\rho(X_M, M, t)}{\partial t} = -\{\rho(X_M, M, t), H(X_M)\} + f(X_M, t) - \hat{Q}\rho(X_M, M, t), \tag{15}$$

where $f(X_M, t) = \sum_{N=0}^{\infty}\int dX_N'[K_{MN}(X_M, X_N')\rho(X_N', t)]$ and $\hat{Q}(*) = \sum_{N=0}^{\infty}\int dX_N'[K_{NM}(X_N', X_M), *]$. This corresponds, in the BL model, i.e., the classical equivalent of Lindbald operator, to the stochastic term, which regulates the jump in state of the system due to the interaction with the reservoir. The definition of a (classical) Lindbald operator (which implies a definition of reservoir and of interaction with the system) has direct consequences in the calculation and physical interpretation of some properties. Here, we report the case of equilibrium time correlation functions. The standard definition of the equilibrium time correlation function at fixed N (e.g. in an NVT ensemble) between two physical observables A and B is [47]:

$$C_{AB}(t) = \frac{1}{Q_N}\int d\mathbf{p}d\mathbf{q}e^{-\frac{H_N(\mathbf{p},\mathbf{q})}{kT}}a(\mathbf{p}, \mathbf{q})b(\mathbf{p}_t(\mathbf{p}, \mathbf{q}), \mathbf{q}_t(\mathbf{p}, \mathbf{q})). \tag{16}$$

Here, $a(\mathbf{p}, \mathbf{q})$ and $b(\mathbf{p}, \mathbf{q})$ are functions in phase space corresponding to A and B, Q_N is the Canonical partition function and $H_N(\mathbf{p}, \mathbf{q})$ the Hamiltonian of a system with N molecules, $\mathbf{p}_t(\mathbf{p}, \mathbf{q})$, $\mathbf{q}_t(\mathbf{p}, \mathbf{q})$ correspond to the evolution in time, at time t, of the momenta and positions with $\mathbf{p}, \mathbf{q}$ initial condition. $C_{AB}(t)$ is determined by calculating $a(\mathbf{p}, \mathbf{q})$ and $b(\mathbf{p}_t(\mathbf{p}, \mathbf{q}), \mathbf{q}_t(\mathbf{p}, \mathbf{q}))$ along the trajectories of the system and then taking the average. In such a case, the dynamics generated is well defined, since the Hamiltonian of N molecules is well defined at any time t:

$$iL = \sum_{j=1}^{N}\left[\frac{\partial H}{\partial \mathbf{p}_j}\frac{\partial}{\partial \mathbf{q}^j} - \frac{\partial H}{\partial \mathbf{q}^j}\frac{\partial}{\partial \mathbf{p}_j}\right] = \{*, H\}\,. \tag{17}$$

However, what happens in case of the GC (μ V, T) ensemble?

Let us start by extending Eq. (17):

$$C_{AB}(t) = \frac{1}{Q_{GC}}\sum_N\int d\mathbf{p}_N d\mathbf{q}_N e^{-\frac{[H_N(\mathbf{p}_N,\mathbf{q}_N)-\mu N]}{kT}}a(\mathbf{p}_N, \mathbf{q}_N)b(\mathbf{p}_t(\mathbf{p}_N, \mathbf{q}_N), \mathbf{q}_t(\mathbf{p}_N, \mathbf{q}_N)), \tag{18}$$

where Q_{GC} is the GC Partition function, μ the chemical potential and N the number of particles. The key question concerns the interpretation of $b(\mathbf{p}_t(\mathbf{p}_N, \mathbf{q}_N), \mathbf{q}_t(\mathbf{p}_N, \mathbf{q}_N))$. In fact, at a given time t the system may have a number of particles N' different from the initial state. In the Kossakowski–Lindblad classical equation, under the condition of flux balance (so that we have the GC ensemble), $b(\mathbf{p}_t(\mathbf{p}_N, \mathbf{q}_N), \mathbf{q}_t(\mathbf{p}_N, \mathbf{q}_N))$ are determined by the action of the propagator with an extended Liouville operator: $e^{i(L_N+L^{Lindblad})t}b(\mathbf{p}, \mathbf{q}) = b(\mathbf{p}_t(\mathbf{p}_N, \mathbf{q}_N), \mathbf{q}_t(\mathbf{p}_N, \mathbf{q}_N))$ with $L^{ext} = L_N + L^{Lindblad}$ (as we have seen in the case of BL model). For time correlation functions, based on single-molecule properties, such as, molecular velocity–velocity time autocorrelation functions or molecular dipole–dipole autocorrelation function, this leads to an ambiguity in the calculation. Differently from the case at fixed N, where "*single molecules*" are clearly defined, for open boundary systems one has to consider three cases:

(i) Molecules that remain in the system at all times for the time window considered.
(ii) Molecules, which are initially in the system but then they leave it (within the time window considered).
(iii) Molecules, which are initially not present in the system but enter from the reservoir (within the time window considered).

At this point, the definition of the action of the reservoir and thus, explicitly or implicitly, the exact definition of $L^{Lindblad}$ becomes crucial. Let us consider, as an example, the BL model of reservoir. For (i), the Lindblad operator applied to a single molecule i is such that $e^{i(L^{Lindblad})t} b_i(\mathbf{p}_i, \mathbf{q}_i) = b_i(\mathbf{p}_i, \mathbf{q}_i)$. In this case, the Lindblad operator (by the definition of its action through K_{MN} and K_{NM} in the BL model) does not act directly on molecule i. This means that in terms of physics, the action of the reservoir does not directly modify the microscopic state of molecule i. It follows that in the definition of $C_{AB}(t)$ only the Liouville's operator L_N is involved. The latter is well defined through the extended Liouville's equation and corresponding Liouville's theorem. For (ii) instead, the action $e^{i(L^{Lindblad})t} b_i(\mathbf{p}_i, \mathbf{q}_i)$ (i.e. according to the action of K_{MN} and K_{NM} in the BL model), has the effect of the action of an annihilation operator which destroys the identity of molecule i when adsorbed by the reservoir. It follows that this action removes its contribution to the correlation function by destroying $b_i(\mathbf{p}_i, \mathbf{q}_i)$ corresponding to the specific molecule (i.e. the molecule does not exist anymore and should be no more counted in the averaging process over all molecules; this is substantially different from considering the correlation to be zero but still counting it in the averaging process). The case (iii) is clear, in fact $a_j(\mathbf{p}_j, \mathbf{q}_j)$ of the molecule j, not present in the system before, is not defined and thus is not counted. The main conclusion, of relevance for example in numerical simulations treated later on, is that once the action of the reservoir is specified, then it follows an unambiguous "*numerical*" procedure for the calculation of $C_{AB}(t)$. From the physical point of view, for an open boundary system with a stochastic reservoir, which "destroys" the molecular identity, a correlation function, linked to a physical process, depends on the locality in space and time of the process. This is because of the finite size and the residence time of molecules in the system. However, there is also the possibility of having a subsystem of a large system and consider the embedding larger part as a particle reservoir. Since the microscopic details of the reservoir in such case are known, the Lindblad operator may be explicitly defined through the microscopic interaction in the system-reservoir term of the Hamiltonian. In such a case, locality in time due to the residence time of a molecule in the system may not play a role. Both interpretations have a physical sense but such a physical sense becomes clear only after we have defined the action of the reservoir. So far, we have talked about particle reservoir without explicitly stating whether the reservoir was infinite or finite. Thus, in the treatment before, we have not made an explicit differentiation between the Grand Ensemble and the GC Ensemble. The latter is the sub-case of the former, in the case of reservoir with an infinite number of particles. Indeed, the case of a reservoir with finite number of particle is of key importance for many systems and will be treated in the next section. As a matter of fact, so far, when introducing a reservoir, we have explicitly or implicitly assumed that the reservoir is large enough, so that we can safely treat this latter only considering its large scale properties. We have also assumed that the open system embedded in the reservoir is large enough, so that the standard laws of statistical mechanics and thermodynamics apply to it. However, a real system is characterized by reservoirs with limited number of particles. Furthermore, the studied systems are often too small to be considered with standard thermodynamics and statistical mechanics. In the next subsection, we treat such two cases.

2.5. Thermodynamics of small open boundary systems

Thermodynamics on small scale at theoretical level was introduced by the work of Hill [51,52]. He later on introduced the term "*nanothermodynamics*", a label for systems that cannot be treated under the hypothesis of thermodynamic limit [15,53,54]. The essence of the problem is that far from the thermodynamic limit quantities as enthalpy, Gibbs free energy and the like, are no more additive. Of particular interest, in our review, is the case of small systems embedded in a large reservoir and systems embedded in a finite reservoir. The work of Schnell and coworkers [55–57] provides a clear presentation of the problem of a small system in a large reservoir and put forward some solutions. The starting point is the basic theory of Hill, that is to consider $\mathcal{N}$ replicas of a small open system embedded in a reservoir of volume V_r, at temperature T and either at constant chemical potential for the ith component μ_i or at a given number of particles $N_{i,r}$. Next, the ensemble formed by the $\mathcal{N}$ replicas (i.e. the total system) can be thought to be large enough so that one can assume the standard laws of thermodynamics to hold for the total system. The Gibbs equation for the $\mathcal{N}$ replicas ensemble can be written as:

$$dE_t = TdS_t - p\mathcal{N}dV + \sum_{i=1,n} \mu_i dN_{i,t} + \left(\frac{\partial E_t}{\partial \mathcal{N}}\right)_{S_t,V,N_{i,t}} d\mathcal{N}, \tag{19}$$

where the label t refers to the total system, E, S, p, V, N refers to the internal energy, entropy, pressure, volume, and number of particles, respectively. The quantity $X \equiv \left(\frac{\partial E_t}{\partial \mathcal{N}}\right)_{S_t,V,N_{i,t}}$ represents essentially the novel aspect in the thermodynamics formalism, this is interpreted as the (reversible) work needed to add one replica to the total system, at constant S_t, V, $N_{i,t}$, so that S_t and $N_{i,t}$ are distributed over one more replica. The integration at constant T, V, μ_i and X of Eq. (19) leads to the equation:

$$E_t = TS_t + \sum_{i=1,n} \mu_i N_{i,t} + X\mathcal{N}. \tag{20}$$

The relation between the variables of the total system and those of the small system is: $E_t = \mathcal{N}\bar{E}$; $S_t = \mathcal{N}S$; $N_{i,t} = \mathcal{N}\bar{N}_i$. Here, the bar denotes the average of a single replica in the GC ensemble, while the entropy S is determined via the probability

distribution in N and E and is the same for each replica. Next, we introduce the variables of the small system in Eq. (20) and obtain:

$$\bar{E} = TS + \sum_{i=1,n} \mu_i \bar{N}_i - \hat{p}V, \tag{21}$$

with $X = -\hat{p}V$, if $p = \hat{p}$ then the system can be described by standard thermodynamics. However, when this situation does not occur then the system has to be considered in the class of *nanothermodynamics*. In such a case the Gibbs–Duhem equation reads as:

$$SdT + \sum_{i=1,n} \bar{N}_i d\mu_i - d(\hat{p}V) + pdV = 0. \tag{22}$$

The key point is that $\hat{p}V$ can be defined through its relation with the GC partition function Ξ, i.e.: $\hat{p}V = k_B T \ln(\Xi)$. The question at this point is, how to derive Ξ for the small system. The authors start from the evidence that for a small system the surface energy becomes important and eventually dominant compared to the interaction energy between particles in the small system. They imagine a subsystem with a generic geometry characterized by a surface, by edges and by corners. Next, they define a measure of the size of the box as $V = L^d$ where d is the dimensionality. It follows that the surface area of the box is proportional to L^{d-1} and can be written as $O \equiv c_s L^{d-1}$, the length of the edges is proportional to L^{d-2} and thus it is written as $L_e \equiv c_e L^{d-2}$ and finally the number of corners defined as c_c. For a sphere, for example, $c_s = (36\pi)^{1/3}$; $c_e = 0$; $c_c = 0$; for a cube we have $c_s = 6$; $c_e = 12$; $c_c = 8$. Then, an effective surface energy for the kth component of the fluid is introduced: $E^s_k(\mu_i, T)$, an equivalent energy for the edges: $E^e_k(\mu_i, T)$, and the corners: $E^c_k(\mu_i, T)$. The one-particle canonical distribution function for the kth component is then: $Q_{1,k} = \frac{V}{\Lambda_k^d} + \frac{O}{\Lambda_k^{d-1}} \exp(\beta E^s_k) + \frac{L_e}{\Lambda_k^{d-2}} \exp(\beta E^e_k) + c_c \exp(\beta E^c_k)$, with Λ_k is the usual thermal de Broglie wavelength for the component k. It follows that the $N_1 \ldots\ldots N_n$-particle canonical distribution function is: $Q_{N_1 \ldots\ldots N_n} = \Pi_{k=1,n} Q_{k,N_k} = \Pi_{k=1,n} \frac{1}{N_k!} Q_{1,k}^{N_K}$, and thus the GC partition function:

$$\Xi = \Pi_{k=1,n} \sum_{N_K=0}^{\infty} \exp(\beta \mu_k N_K) Q_{k,N_k}. \tag{23}$$

Eq. (23) allows us to calculate the size effects due to small size of the system embedded in the reservoir. The major problem with this approach is that it allows only for an estimate of the size effect. In fact, there are two strong approximations involved, on which this approach is based: (a) the neglecting of the interactions between particles in the system; (b) the definition of $E^s_k(\mu_i, T)$, $E^e_k(\mu_i, T)$, $E^c_k(\mu_i, T)$. Rigorous definitions of the boundary energy with a microscopic justification are not known and the authors do not explicitly discuss this aspect. Thus, the model remains in this sense empirical. However, something more specific can actually be said: in essence the boundary energy (that is: $E^b(\mu, T) = E^s(\mu, T) + E^e(\mu, T) + E^c(\mu, T)$ in the formalism used) is the effective energy, which is neglected, when deriving the GC distribution of a system from the canonical distribution of a larger system (where the remaining part is large enough to be considered as a reservoir). This derivation is done in many textbooks of statistical mechanics (see e.g. [58]). Here, we report the part of interest to our discussion. The canonical partition function of a large system is partitioned as: $Z_N(V, T) = \sum_{N_1=0}^{N} \frac{1}{h^{3N_1} N_1! h^{3N_2} N_2!} \int_{\Omega_1} dX_1 e^{-\beta H(X_1)} \int_{\Omega_2} dX_2 e^{-\beta H(X_2)}$, where the number of particles of the total system is partitioned (in all possible combinations) between the system 1 and the system 2: $N = N_1 + N_2$; the volume $V = \Omega_1 + \Omega_2$ and the total Hamiltonian is approximated by $H(X_1, X_2) = H(X_1) + H(X_2)$, with X_1 and X_2 the set of variables in phase space for system 1 and system 2. The term $\mathcal{U}(X_1, X_2)$, which is the direct interaction between system 1 and system 2, is neglected because of the hypothesis usually done that system 1 and system 2 are large enough, so that the energy due to the interactions among only the particle of 1 and only the particle of 2 are dominant. From the equation above the GC partition function can be easily derived. However, for small systems, that is when (for example) 1 is very small, the hypothesis of $\mathcal{U}(X_1, X_2)$ being negligible is no more valid. An estimate of the boundary energy may be thought as the average energy of the ideal boundary between 1 and 2 in the total system: $E^b(\mu, T) = \frac{1}{Z_N(V,T)} \int_V \mathcal{U}(X_1, X_2) e^{-\beta(H(X_1)+H(X_2)+\mathcal{U}(X_1,X_2))}$. The only established rigorous result regarding the surface energy of a large canonical ensemble partitioned in subsystems is the definition of U in the Peierls–Bogoliubov inequality (see e.g. [59,60])

$$Z(N, \Omega) \geq e^{-\beta U} \Pi_{i=1}^{D} Z(N_i, \Omega_i). \tag{24}$$

Here, D is the number of domains, in which the total system is ideally divided, and U is the average interdomain energy in an ensemble, where each domain is independent from the others. Ω is the volume of the total system and Ω_i is the subvolume of the $i-$th domain. For our specific case, consider $i = 1, 2$, where 1 labels the system and 2 the reservoir. However, Eq. (24) refers to a single partitioning of 1 and 2, and $U = \frac{1}{Z_N(V,T)} \int_\Omega \mathcal{U}(X_1, X_2) e^{-\beta(H(X_1)+H(X_2)+\mathcal{U}(X_1,X_2))}$ (i.e. N_1 and N_2 fixed). This implies that it cannot be applied straightforwardly to our case. Nevertheless, it may suggest a rigorous path for the calculation of $E^b(\mu, T)$. (It must be reported that recent results proposed a two-sided Bogoliubov inequality that bounds from above and below the interaction energy of two subsystems [61].) Thus, the approach of Schnell and coworkers is certainly of practical importance. However, in perspective, it should be considered only as a starting point for the treatment of open boundary systems that do not fall under the category of standard thermodynamics. We have suggested that the path

indicated by the Peierls–Bogoliubov inequality can complement the treatment by sharpening a microscopic definition of the corrective/boundary energy. Moreover, the case, where the reservoir is finite, is treated only for an ideal gas, because such a simplification allows for the possibility of a straightforward analytic treatment. Hence, such an approach cannot be sufficient for most of the realistic systems of condensed matter. Instead, regarding the possibility of treating a reservoir with finite number of particles (i.e. the general case of the Grand ensemble), the thermodynamic theory of fluctuations, as presented by Mishin [62], provides a powerful framework.

2.6. Reservoirs with finite number of particles

The starting point of Mishin is the probability density of what he calls internal parameters $X_1, \ldots, X_M$ (i.e. extensive quantities such as volume, energy number of particles):

$$W(X_1, \ldots, X_m) = W_m \exp\left[\frac{S(X_1, \ldots, X_m) - S(X_1^0, \ldots, X_m^0)}{k}\right] \times \left[\frac{S_r(\hat{X}_1 - X_1, \ldots, \hat{X}_m - X_m) - S_r(\hat{X}_1 - X_1^0, \ldots, \hat{X}_m - X_m^0)}{k}\right], \tag{25}$$

where W_m is a normalization factor, $S(X_1, \ldots, X_m)$ the entropy of the system, $X_1^0, \ldots, X_m^0$ the equilibrium values of the quantities considered, k is proportional to the temperature, and $S_r(\hat{X}_1 - X_1, \ldots, \hat{X}_m - X_m)$ is the entropy of the reservoir. The entropy of the reservoir is expanded in Taylor series around $X_1^0, \ldots, X_m^0$ up to the second order: $S_r(\hat{X}_1 - X_1, \ldots, \hat{X}_m - X_m) - S_r(\hat{X}_1 - X_1^0, \ldots, \hat{X}_m - X_m^0) = -\sum_{i=1,m} F_i^0(X_i - X_i^0) + \frac{1}{2}\sum_{i,j=1}^{m}\left(\frac{\partial^2 S_r}{\partial X_i^r \partial X_j^r}\right)_0 (X_i - X_i^0)(X_j - X_j^0)$. Next, he shows that $\left(\frac{\partial^2 S_r}{\partial X_i^r \partial X_j^r}\right)_0$ scales as $1/N_r$ and thus for a large reservoir the second order can be neglected. This hypothesis is named "*reservoir approximation*". In this approximation, he then derives a "*generalized canonical distribution*" starting from the definition of the probability P of a single micro-state. Under the hypothesis of working on a quasi-equilibrium time scale (such that the system can be treated as in equilibrium and considered isolated with fixed value of $X_1, \ldots, X_m$) $S(X_1, \ldots, X_m)$ is the equilibrium entropy. It follows: $W(X_1, \ldots .X_m) = P \ln \Omega_{max}$, where $S(X_1, \ldots, X_m) = k \ln \Omega_{max}$ (being Ω the number of microstates). Next, one has:

$$P = W(X_1, \ldots, X_m) \exp\left[-\frac{S(X_1, \ldots, X_m)}{k}\right], \tag{26}$$

which, under the first order approximation for $W(X_1, \ldots, X_m)$, gives:

$$P = W_m \exp\left[\frac{S(X_1^0, \ldots, X_m^0) - \sum_{i=1}^{m} F_i^0(X_i - X_i^0)}{k}\right]. \tag{27}$$

Eq. (27) can be rewritten as:

$$P = A \exp\left[-\frac{1}{k}\sum_{i=1}^{m} F_i^0 X_i\right], \tag{28}$$

with $A = W_m \exp\left[-\frac{S(X_1^0, \ldots, X_m^0) - \sum_{i=1}^{m} F_i^0 X_i^0}{k}\right]$ being a constant. Eq. (28) is then called the generalized canonical distribution. If, for example, one of its extensive parameter is the energy E, then one can write Eq. (28) as: $P = A \exp\left(-\frac{E + \sum_{2,m} TF_i^0 X_i}{k_B T}\right)$, where T is the temperature. The standard distribution functions are obtained by allowing one or two of the extensive quantities to fluctuate; if E can fluctuate and V and N are fixed, then one has the standard canonical (NVT ensemble), if energy and volume are allowed to fluctuate and N is fixed then one has $X_1 = E$; $X_2 = V$ and $F_2 = p/T$ and thus: $P = A\left(-\frac{E + pV}{k_B T}\right)$, that is the NpT ensemble. Finally, if $X_2 = N$ and $F_2 = -\mu/T$ and V is fixed, then we have the GC ensemble: $P = A\left(-\frac{E - \mu N}{k_B T}\right)$. At this point, if the reservoir is not large enough, so that the term at the second order in the expansion of $S_r(\hat{X}_1 - X_1, \ldots, \hat{X}_m - X_m) - S_r(\hat{X}_1 - X_1^0, \ldots, \hat{X}_m - X_m^0)$ cannot be neglected, then we would have the expression of P with this additional term: $S_r(\hat{X}_1 - X_1, \ldots, \hat{X}_m - X_m) - S_r(\hat{X}_1 - X_1^0, \ldots, \hat{X}_m - X_m^0) = -\sum_{i=1,m} F_i^0(X_i - X_i^0) - \frac{1}{2}k\sum_{i,j=1}^{m} \Lambda_{ij}^r (X_i - X_i^0)(X_j - X_j^0)$, where $\Lambda_{ij}^r = -\frac{1}{k}\left(\frac{\partial^2 S_r}{\partial X_i^r \partial X_j^r}\right)_0$. It follows that:

$$P = A \exp\left[-\frac{1}{k}\sum_{i=1}^{m} F_i^0 X_i - \frac{1}{2}\Lambda_{ij}^r (X_i - X_i^0)(X_j - X_j^0)\right]. \tag{29}$$

For the case with fixed V and fluctuating E and N, from Eq. (29), one finally has:

$$P = A \exp\left[\frac{E - \mu N}{k_B T} - \frac{1}{2}\Lambda_{22}^r (N - N_0)^2\right], \tag{30}$$

where the elements of reservoir matrix Λ_{ij} are such that: $\lambda_{11}, \Lambda_{12}, \Lambda_{21} = 0 \;\; ; \Lambda_{22} = \rho_0 c_v T(\partial\mu/\partial\rho)_T + \mu_0$. Here, the subscript 0 indicates the average/equilibrium value, c_v is the heat capacity at constant volume ρ the density. The fact that all elements of the matrix are zero, except Λ_{22}, is due to the fact that we treat the reservoir as a finite source of particles but an infinite source of heat. Thus, $\Lambda_{11} = 0$, while for simplicity it is assumed that the chemical potential μ is independent of the temperature. Since $\Lambda_{12} = \Lambda_{21}$, we have that the off-diagonal terms are also zero. Eq. (30) can be reduced to an expression similar to that of the usual GC ensemble:

$$P = A \exp\left[\frac{E - \hat{\mu}N}{k_B T}\right], \tag{31}$$

where $\hat{\mu} = \mu_0 + \Lambda^r_{22} k_B T N_0 - \frac{1}{2}\Lambda^r_{22} k_B T N$. Thus, instead of a fixed chemical potential we have an effective chemical potential depending in a linear way from the instantaneous number of particles N. If $N_r \to \infty$ then $\Lambda^r_{22} \to 0$ and one recovers the standard GC distribution. In practical applications, the effect of the finite reservoir can be applied in a Monte Carlo (MC) scheme by considering an adjustable chemical potential on-the-fly, see, e.g., Refs. [63–65]. The theoretical landscape about open boundary systems outlined in this section, can be now used as a reference for computational implementations and applications or as a basis for the development of novel algorithms for future applications. In the next section, we will focus on the implementation of the open boundary system idea in molecular simulation. We will point, as often as possible, on the consistency and inconsistency with the theoretical work and suggest, when possible, potential connections for mutual improvements.

3. Molecular simulation of open systems

In this section, our attention will be devoted to molecular systems in thermodynamical equilibrium. Molecular simulations of open systems with fluctuating number of particles are then, naturally, closely related to simulations in the GC statistical ensemble. The latter are typically carried out using the MC approach (see Appendix B) because the associated insertion and deletion of particles can be relatively easily incorporated in the stochastic MC moves [66]. This is more difficult to accomplish in the case of Molecular Dynamics (MD) (see Appendix A) and consequently the number of numerical methods for carrying out MD simulations in the GC ensemble is much lower. However, it is precisely the insertion and deletion of particle that gives motivation for the development of GC MD methods. Namely, the probability of particle insertion is related to a random density fluctuation that generates a cavity, where we can insert a particle. As the probability for such an event is rather low in dense systems, the success rate of insertion also becomes low, which makes GC MC simulations prohibitively expensive for many such molecular systems. A similar problem of particle insertion one encounters also in computations of chemical potential [67] as well as open boundary molecular simulations, discussed in Section 3.2.2 later on, and in methods that couple atomistic and continuum hydrodynamics (see Section 3.2.3). There, we also report our solution to circumvent it. Since examples of MC simulations in GC ensemble can be found in many textbooks, e.g., Refs. [47,66,68], we will not delve into them here. Rather, we briefly report some examples of MD methods for simulations in the GC ensemble published in the literature. Pettitt et al. [69–73] have developed MD methods for GC simulations based on extended phase space approach, where a classical system is coupled to a chemical potential reservoir with an additional variable, resembling extended phase space approaches for performing MD in canonical ensemble of at constant pressure [47]. The additional variable governs the dynamics of the variation of number of particles in the system. The integer part of the continuous variable stands for the number of full particles while the fractional part represents a fractional particle whose coupling to the physical system is scaled. MD methods based on similar ideas of extended Hamiltonian formalism are also reported in Refs. [74–78]. Another branch of this type of methods is represented by hybrid MD–MC methods using MC variations of the continuous coupling variable [12,79,80]. All the mentioned methods have been, thus far, successfully applied to rather simple molecular systems such as ideal gas, Lennard–Jones fluid at different phase points, and liquid water. Their application to more complex systems is hindered by the fact that only one fractional particle can be inserted per timestep leading to slow equilibration and a tendency to get stuck in metastable states for a long time [77]. A solution for the problem of efficient insertion of particles into dense liquids, which opens doors for open MD simulations of complex fluids, e.g., star-polymers melt [81,82], is provided by adaptive resolution simulation, discussed next.

3.1. Adaptive resolution simulation

Single-scale (all-atom) MD and MC methods, discussed above, are extremely valuable simulation tools as they can provide detailed insight into the structural and dynamical properties of complex soft and biological matter systems [83]. Thus, they allow us to understand the microscopic origin of macroscopic properties. Unfortunately, performing such simulations still represents a computational challenge, even for the cutting edge computers, because of the inability to cover all time and length scales associated with physical phenomena of interest. Accessing wide range of spatiotemporal scales thus inevitably demands some simplifications (coarse-graining) of the molecular model with a large number of DOFs. Typically, in biomolecular simulations this high computational cost is associated with simulation of the solvent. Water, the most abundant solvent in nature, is of essential importance for proper functioning and stability of (bio)macromolecules. However, it often comprises most of the particles in the simulated system and hence the majority of the computational effort is spent computing interactions between solvent molecules far away from the macromolecule, in distant domains that are not relevant for the problem under consideration. This realization led to the development of several multiscale simulation methods, which reduce the number of DOFs for distal water and at the same time keep the high (atomistic) resolution where it is necessary.

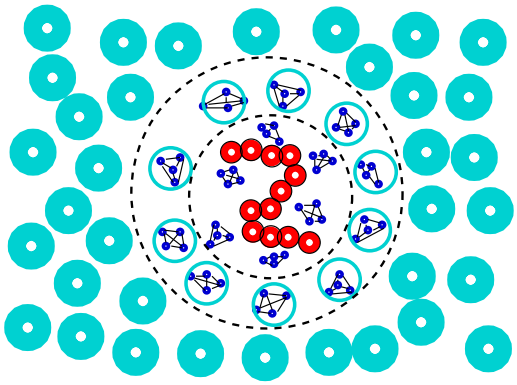

Fig. 1. The basic idea of adaptive resolution: a cartoon of a solvated generic macromolecule. The proximal solvent within a certain radius from the macromolecule is represented using a fine-grained resolution. The distal solvent molecules are represented by a coarse-grained model as single beads. *Source:* Reprinted from Ref. [89].

In general, multiscale simulation methods can be classified into two groups: hierarchical (sequential or homogenization) and concurrent (domain decomposition) methods [84,85]. In the first type of methods, the whole computational domain is treated on one level of resolution at a time. The information obtained at a finer resolution is used to parametrize a lower-resolution model. Back mapping methods then enable us to go back and forth between the two resolutions by reintroducing the chemical details on demand [86–88]. On the other hand, in multiscale methods of the second type, multiple levels of resolution are applied in simulation box concurrently. A typical example, presented in Fig. 1, is a solvated macromolecule where the interesting physics occurs within a few solvation shells around the macromolecule, while the distal solvent plays a role of a buffer and may be therefore treated on a coarser level. The link between different levels of resolution can be done either by a fixed-resolution approach where different resolution domains interact with each other but do not exchange molecules, e.g., QM/MM approach (see Section 4.2), or by the adaptive resolution approach where molecules change their resolution according to their location in the simulation box.

In the following, we will focus on the concurrent multiscale methods. We will first provide a review of the basic theoretical principles of adaptive resolutions scheme (AdResS) [85,90–92]. AdResS allows for a true equilibrium between high and low resolution domains of a system within an MD simulation. The method can link two or more resolution regions that span from quantum all the way to continuum length scales of soft matter systems. Next, we will review also its Hamiltonian version, i.e., H-AdResS, and address their similarities and differences, followed by a comparison with other existing concurrent multiscale molecular methods. But first, we will shed some light on the connection between adaptive resolution methods and alchemical free energy perturbation methods [93]. The similarity may not be that surprising as the change of resolution has some resemblance to alchemical transformation between the fine- and coarse-grained models [90].

3.1.1. Alchemical free energy perturbation methods: thermodynamic perturbation and integration

Concurrent multiscale simulations, where two different models are in thermodynamic coexistence within the same simulation box, are closely related to free-energy perturbation methods, where one transforms a system with N particles from one thermodynamic state A to another state B [47,68,93–95]. The corresponding Hamiltonians are $H^A(\mathbf{p}^N, \mathbf{r}^N) = \sum_{i=1}^N \mathbf{p}_i^2/2m_i + U^A(\mathbf{r}^N)$ and $H^B(\mathbf{p}^N, \mathbf{r}^N) = \sum_{i=1}^N \mathbf{p}_i^2/2m_i + U^B(\mathbf{r}^N)$. The free energy difference between the two systems can be then computed using the free energy perturbation formula as [94]

$$\begin{aligned} \Delta F^{AB} &= -k_B T \ln \left\langle \exp\left[-\frac{H^B(\mathbf{p}^N, \mathbf{r}^N) - H^A(\mathbf{p}^N, \mathbf{r}^N)}{k_B T} \right] \right\rangle_A \\ &= -k_B T \ln \left\langle \exp\left[-\frac{U^B(\mathbf{r}^N) - U^A(\mathbf{r}^N)}{k_B T} \right] \right\rangle_A , \end{aligned} \tag{32}$$

where $\langle \cdots \rangle_A$ denotes averaging over the ensemble of configurations representative of the thermodynamic state A.

The above approach becomes inaccurate if the states A and B do not have overlap in configuration space. Then $|U^B - U^A| \gg k_B T$ and $\exp(-(U^B - U^A)/k_B T)$ becomes very small. Consequently, we do not sample adequately the configuration space of B when performing the averaging $\langle \cdots \rangle_A$. To circumvent this problem, the transformation between the states A and B is replaced by a series of intermediate states along a pathway that connects the states A and B. The pathway is parameterized with a coupling parameter $\lambda \in [0, 1]$. As the coupling parameter λ changes from 0 to 1, the Hamiltonian varies from H^A to H^B. Hamiltonians of the intermediate states are $H^\lambda(\mathbf{p}^N, \mathbf{r}^N) = \sum_{i=1}^N \mathbf{p}_i^2/2m_i + U^\lambda(\mathbf{r}^N)$ with

$$U^\lambda(\mathbf{r}^N) = (1-\lambda)U^A(\mathbf{r}^N) + \lambda U^B(\mathbf{r}^N) = \lambda U^B(\mathbf{r}^N) + (1-\lambda)U^A(\mathbf{r}^N). \tag{33}$$

The free energy difference ΔF_{AB} is the sum of free-energy differences from all the intermediate states along the pathway

$$\begin{aligned}\Delta F^{AB} &= -k_BT\sum_{k=1}^{M-1}\ln\left\langle\exp\left[-\frac{H^{\lambda_{k+1}}(\mathbf{p}^N,\mathbf{r}^N)-H^{\lambda_k}(\mathbf{p}^N,\mathbf{r}^N)}{k_BT}\right]\right\rangle_k\\ &= -k_BT\sum_{k=1}^{M-1}\ln\left\langle\exp\left[-\frac{U^{\lambda_{k+1}}(\mathbf{r}^N)-U^{\lambda_k}(\mathbf{r}^N)}{k_BT}\right]\right\rangle_k,\end{aligned} \tag{34}$$

where index k runs over the initial and final states A and B, respectively, and all $M-2$ intermediate states with $\lambda_1=0$ and $\lambda_M=1$, so that $k=1$ and $k=M$ correspond to the states A and B, respectively.

In practical applications, when we alchemically transform a molecular species into another one, we can distinguish three different groups of particles in the system: a group of particles that do not change during the simulation run, a group that contains the particles of the initial state A, and a group of particles corresponding to the final state B. The hybrid Hamiltonian of the system can be written as [96]

$$H^\lambda = H^0 + (1-\lambda)H^a + \lambda H^b, \tag{35}$$

where H^0 is the part of the Hamiltonian corresponding to the particles that remain unchanged during the simulation, H^a and H^b the parts corresponding to atoms characterizing the initial and final states A and B, respectively.

Another method for free energy difference computation is the thermodynamic integration [47,68,93]. The free energy difference ΔF^{AB} is calculated as

$$\Delta F^{AB} = \int_0^1 \frac{\partial F}{\partial\lambda}\,d\lambda = \int_0^1\left\langle\frac{\partial H^\lambda(\mathbf{p}^N,\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda d\lambda = \int_0^1\left\langle\frac{\partial U^\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda d\lambda = \int_0^1\left\langle U^B-U^A\right\rangle_\lambda d\lambda, \tag{36}$$

where $\langle\cdots\rangle_\lambda$ denotes averaging over the ensemble of configurations representative of the intermediate state characterized with a fixed value of the coupling parameter λ.

In practice, M molecular simulations at M discrete values of the coupling parameter $\lambda\in[0,1]$ have to be conducted. For each value of λ, the average $\left(\frac{\partial U^\lambda(\mathbf{r}^N)}{\partial\lambda}\right)_\lambda$ is computed, yielding [93]

$$\Delta F^{AB} = \sum_{k=1}^{M}\left\langle\frac{\partial U^\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_{\lambda_k}\Delta\lambda_k, \tag{37}$$

where $\Delta\lambda_k=\lambda_{k+1}-\lambda_k$.

3.1.2. Adaptive resolution scheme (AdResS)

Let us now consider the situation depicted in Fig. 1 and couple the fine- and coarse-grained solvent models. The two molecular models should be in thermodynamic equilibrium within the same simulation box. Furthermore, we would like the solvent molecules to freely move across different resolution domains and change their resolution on the fly accordingly. To this end, let us exploit the analogy with alchemical transformations presented in the previous subsection. We can imagine that when molecules cross the resolution domains they are being alchemically transformed from one species to another.

The switching of the resolution can then be described by making the coupling parameter λ position dependent so that the different values of λ would determine the level of resolution according to the position in the simulation box [90]. For simplicity, let us only consider two-body interactions, but our consideration can be generalized to many-body cases as well. We can then set $\lambda_{\alpha\beta}=\lambda_{\alpha\beta}(\mathbf{R}_\alpha,\mathbf{R}_\beta)\in[0,1]$, where $\mathbf{R}_\alpha$ and $\mathbf{R}_\beta$ are the center-of-mass (CoM) positions of two given molecules α and β, respectively. Inspired by Eq. (33), the coupling of the coarse-grained model A and fine-grained model B can be done as

$$U^\lambda = \frac{1}{2}\sum_\alpha\sum_{\beta\neq\alpha}U^\lambda_{\alpha\beta} = \frac{1}{2}\sum_\alpha\sum_\beta\left[\lambda_{\alpha\beta}U^B_{\alpha\beta}+(1-\lambda_{\alpha\beta})U^A_{\alpha\beta}\right]. \tag{38}$$

Next, we need to define the functional form of $\lambda_{\alpha\beta}$. As the simplified coarse-grained model A has less DOFs than the fine-grained model B, we can make a natural choice and require that a coarse-grained molecule can interact with any other molecule (even in the high resolution domain) only through coarse-grained effective potential. A simple choice that satisfies this condition is $\lambda_{\alpha\beta}=\lambda_{\alpha\beta}(\mathbf{R}_\alpha,\mathbf{R}_\beta)=w(R_\alpha)w(R_\beta)=w_\alpha w_\beta$, where $w_\alpha=w(R_\alpha)$, $w_\beta=w(R_\beta)$, and $R_\alpha=|\mathbf{R}_\alpha-\mathbf{R}|$ and $R_\beta=|\mathbf{R}_\beta-\mathbf{R}|$ are distances of solvent molecules α and β from the CoM of the macromolecule $\mathbf{R}$, respectively. The weighting function w is a smooth continuous function that interpolates between values 0 and 1, where the former defines the region of model A and the latter the domain belonging to model B. The values $0<w<1$ correspond to the transition region sandwiched in between the former two. The purpose of the transition region is that the introduction of the fine-grained DOFs, when a given molecule enters from the coarse-grained side, is continuous and not instantaneous leading to overlaps with the atoms of the neighboring molecules. Thus, the transition region enables the molecule to gradually find an energetically

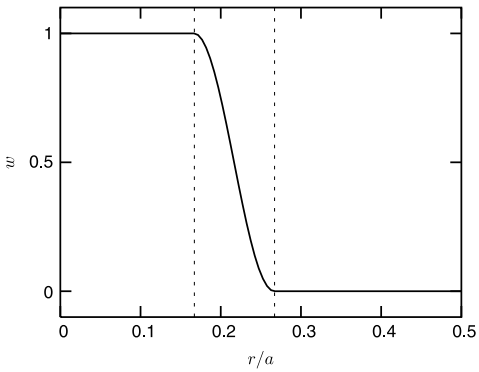

Fig. 2. The weighting function $w(r) \in [0, 1]$ defined by Eq. (39). The values $w = 1$ and $w = 0$ correspond to the fine-grained region containing molecules of the model B and coarse-grained region of the model A, respectively inside of the simulation box a. The vertical lines denote the boundaries of the transition region.
Source: Reprinted from Ref. [97].

permissible orientation with respect to its neighboring molecules. A simple choice for w, corresponding to the spherical geometry in Fig. 1, is [97]

$$w(r) = \begin{cases} 1; & 0 \le r < r_0 \\ \cos^2\left[\frac{\pi}{2d}(r - r_0)\right]; & r_0 \le r < r_0 + d \\ 0; & r_0 + d \le r \end{cases} \tag{39}$$

where r_0 is the radius of the fine-grained region B around the CoM of the macromolecule and d the width of the transition region. The weighting function w is depicted in Fig. 2. Using $\lambda_{\alpha\beta}$ defined in this way, we can express $U_{\alpha\beta}$ in Eq. (38), dropping the superscript λ, as [98]

$$U_{\alpha\beta} = \lambda_{\alpha\beta} U^B_{\alpha\beta} + (1 - \lambda_{\alpha\beta}) U^A_{\alpha\beta} = w_\alpha w_\beta U^B_{\alpha\beta} + (1 - w_\alpha w_\beta) U^A_{\alpha\beta}. \tag{40}$$

The corresponding conservative force acting on the molecule α is [99]

$$\mathbf{F}_{\alpha\beta} = -\frac{\partial U_{\alpha\beta}}{\partial \mathbf{R}_\alpha} = w_\alpha w_\beta \mathbf{F}^B_{\alpha\beta} + (1 - w_\alpha w_\beta)\mathbf{F}^A_{\alpha\beta} + \frac{\partial w_\alpha}{\partial \mathbf{R}_\alpha} w_\beta \left[U^A_{\alpha\beta} - U^B_{\alpha\beta}\right], \tag{41}$$

where $\mathbf{F}^A_{\alpha\beta} = -\frac{\partial U^A_{\alpha\beta}}{\partial \mathbf{R}_\alpha}$ and $\mathbf{F}^B_{\alpha\beta} = -\frac{\partial U^B_{\alpha\beta}}{\partial \mathbf{R}_\alpha}$.

The respective force acting on the molecule β is

$$\mathbf{F}_{\beta\alpha} = -\frac{\partial U_{\alpha\beta}}{\partial \mathbf{R}_\beta} = w_\alpha w_\beta \mathbf{F}^B_{\beta\alpha} + (1 - w_\alpha w_\beta)\mathbf{F}^A_{\beta\alpha} + \frac{\partial w_\beta}{\partial \mathbf{R}_\beta} w_\alpha \left[U^A_{\alpha\beta} - U^B_{\alpha\beta}\right]. \tag{42}$$

Using $\mathbf{F}^A_{\beta\alpha} = -\mathbf{F}^A_{\alpha\beta}$ and $\mathbf{F}^B_{\beta\alpha} = -\mathbf{F}^B_{\alpha\beta}$ we obtain

$$-\mathbf{F}_{\beta\alpha} = w_\alpha w_\beta \mathbf{F}^B_{\alpha\beta} + (1 - w_\alpha w_\beta)\mathbf{F}^A_{\alpha\beta} - \frac{\partial w_\beta}{\partial \mathbf{R}_\beta} w_\alpha \left[U^A_{\alpha\beta} - U^B_{\alpha\beta}\right]. \tag{43}$$

Should the force given by Eq. (41) satisfy Newton's Third Law, i.e., $\mathbf{F}_{\alpha\beta} = -\mathbf{F}_{\beta\alpha}$, from above equations it follows that this force, despite being conservative, can satisfy Newton's Third Law only for a trivial case of $w = const.$, corresponding to constant resolution simulations. Otherwise, we are left with the third term, i.e., a drift force

$$\mathbf{F}^{dr}_{\alpha\beta} = \frac{\partial w_\alpha}{\partial \mathbf{R}_\alpha} w_\beta \left[U^A_{\alpha\beta} - U^B_{\alpha\beta}\right], \tag{44}$$

which violates Newton's Third Law [98–102]. This drift force involves derivatives of the weighting function w and hence acts only in the transition region, where the switching of resolution takes place. Due to the drift force linear momentum cannot be conserved using Eq. (40). This is to be expected as the translational invariance is broken due to the resolution change. Another point to consider is that potential energies $U^A_{\alpha\beta}$ and $U^B_{\alpha\beta}$ are determined up to some constants. This does not effect the dynamics in respective monoscale simulations as forces depend on gradients of the potentials. However, here they enter as a linear term in the drift force determining its size. Despite this conceptual limitation, from a technical point of view Eq. (40) can be used to develop a variant of AdResS with a Hamiltonian formulation (H-AdResS) [103,104] (see Section 3.1.3).

Instead, to satisfy Newton's Third Law and strictly conserve the total linear momentum of the system, which is crucial for hydrodynamics, we couple the two molecular models A and B with the force coupling scheme, omitting the drift force $\mathbf{F}^{dr}_{\alpha\beta}$:

$$\mathbf{F}_{\alpha\beta} = \lambda_{\alpha\beta}\mathbf{F}^{B}_{\alpha\beta} + (1-\lambda_{\alpha\beta})\mathbf{F}^{A}_{\alpha\beta} = w_\alpha w_\beta \mathbf{F}^{B}_{\alpha\beta} + (1 - w_\alpha w_\beta)\mathbf{F}^{A}_{\alpha\beta}. \tag{45}$$

Eq. (45) is at the heart of AdResS [85,90]. Note that dynamics governed by the AdResS force given by Eq. (45) is linear momentum conserving despite broken translational invariance, as AdResS is not a Hamiltonian method. But, it will be shown in Section 3.2.1 that for the calculation of statistical and dynamical properties of interest a global Hamiltonian is actually not needed. For the molecular system presented in Fig. 1, the intermolecular force

$$\mathbf{F}^{B}_{\alpha\beta} = \sum_{i\alpha, j\beta} \mathbf{F}^{B}_{i\alpha j\beta} \tag{46}$$

is the sum of all pair atom interactions between explicit atoms of the molecule α and explicit atoms of the molecule β

$$\mathbf{F}^{B}_{i\alpha j\beta} = -\frac{\partial U^B}{\partial \mathbf{r}_{i\alpha j\beta}}. \tag{47}$$

The vector $\mathbf{r}_{i\alpha j\beta} = \mathbf{r}_{i\alpha} - \mathbf{r}_{j\beta}$ is the relative position vector of atoms $i\alpha$ and $j\beta$. The effective coarse-grained force between molecules α and β is

$$\mathbf{F}^{A}_{\alpha\beta} = -\frac{\partial U^A}{\partial \mathbf{R}_{\alpha\beta}}, \tag{48}$$

where $\mathbf{R}_{\alpha\beta} = \mathbf{R}_\alpha - \mathbf{R}_\beta$ is the relative position vector of the CoMs of the molecules α and β, respectively.

The intermolecular potential of the coarser model A, U^A, can be derived in several ways. One way is to map the structure of the coarser model A as closely as possible to the reference fine-grained model B. This is desirable because we want the molecules to adapt, when they enter the high resolution domain from the low resolution region (through the transition one), as quickly as possible to the new environment. (However, as we will see in a bit, this restriction can be substantially relaxed.) To this end, we reduce the potential of the fine-grained model B into a reduced effective potential U^A using any of the standard bottom-up structure-based coarse-graining methods [105,106], e.g., Iterative Boltzmann inversion [107–110], inverse MC [111,112], force-matching scheme (the multiscale coarse-graining method) [113–116], Extended ensemble approach [117,118], relative entropy [119–121] etc.

Note that in the free-energy perturbation approaches introduced previously, the models A and B have the same number of DOFs and the kinetic energy part of Hamiltonian is the same in both models. This is also true for H-AdResS presented later on. Here, as already mentioned, if the model A represents a coarse-grained version of the model B, in general, number of DOFs is lower in the model A (of course, it can be also the same). Thus, each time a molecule leaves (or enters) the region A it gradually gains (or loses) its e.g. vibrational and rotational DOFs explicitly present in the model B while retaining its linear momentum. The switching on/off of these DOFs in the transition region is related to non-integer DOFs, which can be described by fractional calculus [98,99,122–126]. Here, we will not dwell upon it but let us just mention that this has some similarities with the alchemical transformation described by Eq. (35). The coarse-grained part of the Hamiltonian can be considered as H^0, while the respective parts of Hamiltonian associated with the switched off/on DOFs can be related to the H^a and H^b[99]. Furthermore, in the lower resolution domain A each molecule represents many orientations and conformations of the corresponding molecule in the high resolution domain B. Hence, while switching between the representations in the transition regime, one does not exactly reproduce the fine-grained coordinates and velocities. The reverse-mapping thus destroys time-reversibility in the simulation. Because time reversibility is essential for energy conservation [127], AdResS does not conserve energy. In particular, the force in Eq. (45) is in general not conservative in the transition region (i.e., in general $\oint \mathbf{F}_{ab} \cdot \mathbf{dr} \neq 0$) [99–101] and global potential energy along the lines of Eq. (38) and corresponding global Hamiltonian (sum of kinetic and potential energies) are not defined. Hence, the conservation of the linear momentum comes at the cost of limiting ourselves to the MD methodology leaving out the possibility of running MC simulations that require defined global potential energy of the system. Moreover, AdResS cannot be run in the microcanonical ensemble and a local thermostat, e.g., Langevin thermostat [128], is needed to supply or remove the free-energy associated with the switched on/off DOFs at the resolution transition [85,91,92]. We will return to this point shortly. Nevertheless, for tackling problems out of thermodynamic equilibrium, e.g., fluid flows, where statistical mechanics is not applicable, and other conserved properties, e.g. linear momentum conservation in hydrodynamics, play a more important role, AdResS based on the force-interpolation ansatz seems to be the method of choice.

For conducting AdResS simulations, we need thermodynamic coexistence of models A and B. In general, the necessary condition for thermodynamical equilibrium between the coarse- and fine-grained models and the transition region, i.e., the chemical potential, pressure, and temperature equivalence, cannot be assured solely by a derivation of the U^A based on U^B. Moreover, we would like to be in a flexible position and couple two models where the coarser model A is not derived in a bottom-up way from the fine-grained model B but can be obtained using a top-down coarse-graining approach, instead. Or, we would also like to consider situations where we couple different molecular models at the same level of resolution and number of DOFs but different force fields. In short, we would like to have a concurrent multiscale approach that allows us to

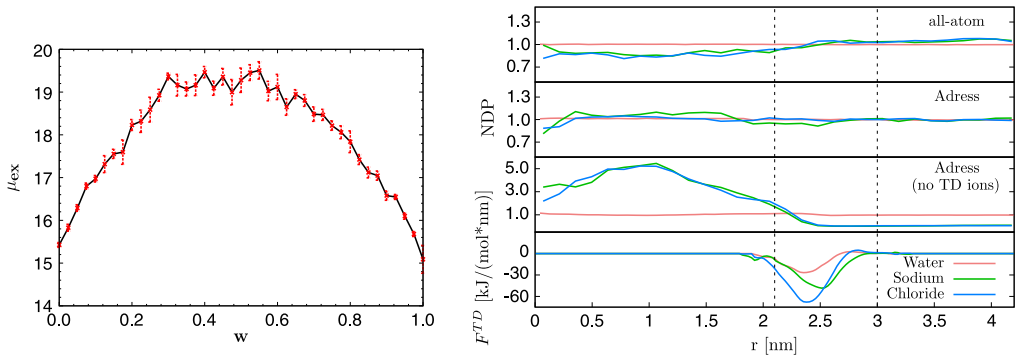

Fig. 3. (left) The excess chemical potential through the transition region. (right) Normalized density profiles for water molecules (red) and sodium (green) and chloride (blue) ions. The results are shown for the full-blown atomistic simulation and AdResS simulations where the thermodynamics forces act on all molecular species or only on water molecules for comparison. The bottom plot shows the TD forces applied to all three molecule types. The thermodynamic forces clearly make the density profiles flat.
Source: Reprinted from Refs. [129,130].

couple rather loosely connected molecular representations, i.e., it maintains two different representations with, in general, different chemical potentials ($\mu_A \neq \mu_B$) in thermodynamic equilibrium.

To treat these scenarios, we extend the original scheme, Eq. (45), by subtracting a thermodynamic force $\mathbf{F}^{TD}$. The total force on molecule α is

$$\mathbf{F}_\alpha^{AdResS} = \sum_{\beta\neq\alpha}\left[w_\alpha w_\beta \mathbf{F}^B_{\alpha\beta} + (1 - w_\alpha w_\beta)\mathbf{F}^A_{\alpha\beta}\right] - \mathbf{F}^{TD}_\alpha, \tag{49}$$

where $\mathbf{F}^{TD}_\alpha = \mathbf{F}^{TD}(R_\alpha)$ is the thermodynamic force defined as a negative gradient of the excess chemical potential μ^{exc} due to the intermolecular interactions [129]. The aim of $\mathbf{F}^{TD}$ is to compensate the chemical potential difference across the simulation box and obtain a flat density profile across the whole system. As mentioned, AdResS is a non-conservative scheme and hence the potential energy is not defined. Therefore, to calculate numerically the excess chemical potential we proceed as follows. We divide the simulation box into regions of force-fields A and B and the transition region in between. The region A is characterized by the value of the switching function $w_0 = 0$. The region B is characterized by the value of the switching function $w_{N+1} = 1$. In the transition region the value of w in the actual simulations varies continuously. Here, we approximate this by discretizing w into N steps $w_1, w_2, \ldots, w_{N-1}, w_N$. For any fixed value of w,[1] the energy function is well defined ($\mathbf{F}_\alpha^{AdResS}$ becomes conservative with a well defined global potential). The excess chemical potential is then defined as: $\mu^{exc}(r_i) = \mu^{exc}_{w_i}$, where the $\mu^{exc}_{w_i}$ is the chemical potential of the molecules in a bulk system of the specific representation of w_i. To calculate numerically each $\mu^{exc}(w_i)$ one can use standard particle insertion methods, e.g., Widom insertion method [67]. Repeating this procedure with all values of w_i leads to a position dependent excess chemical potential $\mu^{exc}(r)$.

An example of $\mu^{exc}(w_i)$ for a liquid of tetrahedral molecules [129] is presented in Fig. 3 (left). The system is set up in such a way that the equation of state is the same in both the fine- and coarse-grained regimes at the temperature and density of the simulation. Because of that μ^{exc} is the same for $w = 1$ and for $w = 0$ but different for $0 < w < 1$. This leads to a nonzero $\mathbf{F}^{TD}$ in the transition region by computing a gradient of the μ^{exc} curve.

As the above procedure of computing $\mathbf{F}^{TD}$ involves several computations of chemical potential, $\mathbf{F}^{TD}$ can be more conveniently cast as the force on a molecule balancing the pressure gradient[2] $-\nabla p(r)$

$$\mathbf{F}^{TD}_\alpha = \frac{M_\alpha}{\rho_0}\nabla p, \tag{50}$$

where M_α and ρ_0 are the mass of molecule α and the reference bulk density, respectively [133]. Numerically, this translates into an iterative procedure [133]

$$\mathbf{F}^{TD}_{\alpha\ i+1} = \mathbf{F}^{TD}_{\alpha\ i} - \frac{M_\alpha}{\rho_0^2 \kappa_T}\nabla\rho_i(r), \tag{51}$$

[1] This is the value that we obtain by using the insertion methods in a hybrid system exclusively composed of hybrid molecules with a fixed level of resolution $0 \leq w = w(r) = const. \leq 1$ corresponding to a fixed bulk value $\mu_{w(r)}$.

[2] Note that we switch from the molecular (particle-based) to field description in the derivation of $\mathbf{F}^{TD}$ (which acts on the centers of mass of molecules). For example, center of mass density field is defined as $\rho(\mathbf{r}) = M_\alpha \sum_\alpha \delta(\mathbf{r} - \mathbf{R}_\alpha)$ and similarly one defines also the pressure and chemical potential density fields, see, e.g., Refs. [131,132]. Therefore, although $\mathbf{F}^{TD}_\alpha$ is not pairwise, it conserves local linear momentum on this coarser level of description. Besides, $\mathbf{F}^{TD}_\alpha = 0$ in the region of interest, i.e., in the fine-grained region B.

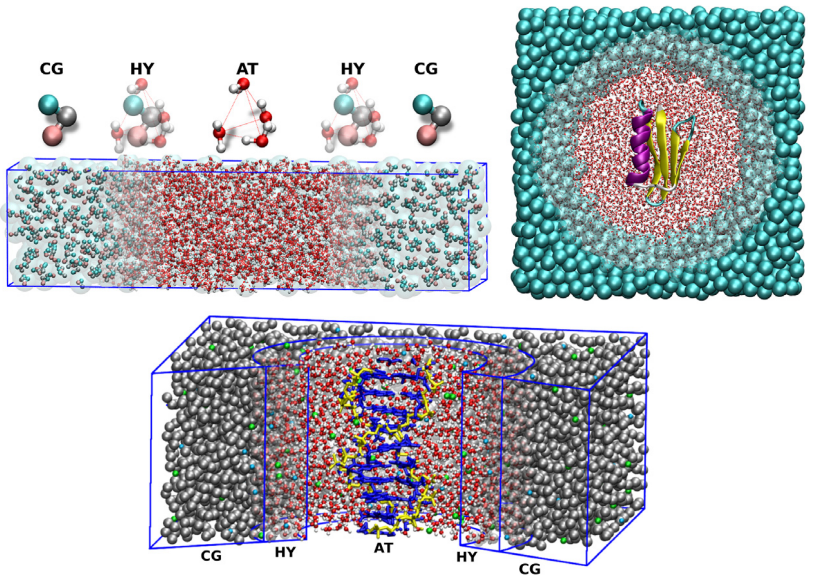

Fig. 4. Adaptive resolution simulations of: (above, left) supramolecular polarizable water with the flat boundary; (above, right) an atomistic protein in supramolecular water with the spherical boundary; (below) a DNA molecule in salt solution with the cylindrical boundary between regions *A* and *B*. Abbreviations AT, HY, CG stand for fully atomistic, hybrid, and coarse-grained regions, respectively.
Source: Reprinted from Refs. [130,141,143].

where κ_T is isothermal compressibility, respectively. The iteration is performed until the system obtains the target uniform density. In practice, we actually use the formula [91]

$$\mathbf{F}^{TD}_{\alpha\ i+1} = \mathbf{F}^{TD}_{\alpha\ i} - C_i \nabla \rho_i(r), \tag{52}$$

where C is a numerical prefactor that is determined empirically. This is similar to other methods for enforcing uniform density profile [134,135]. The value of the prefactor is adjusted along the process to prevent overcorrection. To speedup the iteration procedure we also run simultaneously at each step several simulations with different prefactors and chose the best one for the next iteration. Furthermore, when different types of particles are present in the system, the iteration procedure is applied for all types, as depicted in Fig. 3 (right) for a system of salt solution (1M NaCl) [130,136]. On the other hand, if the *A* model is chosen to be equal to the *B* model (this corresponds to a homogeneous monoscale system) then $\mathbf{F}^{TD}_{\alpha} = 0$ and Eq. (49) simplifies to $\mathbf{F}^{AdResS}_{\alpha} = \sum_{\beta\neq\alpha}\mathbf{F}^{B}_{\alpha\beta} = \mathbf{F}^{B}_{\alpha} = \mathbf{F}^{A}_{\alpha}$.

AdResS defined by Eq. (49) offers ample options: we can achieve model coexistence between models *A* and *B* that are either of the same resolution but use different force fields (e.g., GROMOS [137] or AMBER [138]), coarse- and fine-grained models, which can be structurally and/or thermodynamically mapped on each other or not [133,139], supramolecular couplings [140–145], where one coarse-grained bead corresponds to many fine-grained molecules, e.g., the MARTINI model [146–148], or even coupling atomistic water with even more simplified coarse-grained models like ideal gas [149] or bridging to the hydrodynamics scale as for example in coupling of MD with multiparticle collision dynamics [150] or smoothed dissipative particle dynamics [151]. We can also couple classical and quantum mechanics using Path Integral formalism as described later on in Section 4. The method can accommodate various geometric boundaries between the resolution regions. In examples presented in Fig. 4, we demonstrate domain decomposition with flat, cylindrical, and spherical boundaries. The geometrical boundary between resolution regions can thus be set to reflect the shape of the simulated molecule, i.e., the cylindrical for a DNA molecule and spherical for a protein. The center of the fine-grained region can either be a fixed point (usually the center of the simulation box) or a mobile point, as for example, in a simulation of a macromolecule where it coincides with the macromolecule's CoM. A setup, where the center of the fine-grained region follows the macromolecule's random translation, ensures that the macromolecule is always expressed in high resolution and surrounded by a layer of high-resolution solvent [141,152,153]. Moreover, the boundaries between different domains can be also flexible, changing shape according to a given situation during the course of simulation [154,155]. Furthermore, one can also consider situations where a macromolecule spreads over several resolution domains described with different levels of detail [91].

Coming back to the role of the local thermostat in AdResS simulations, apart from supplying or removing the free energy (latent heat) associated with the switching of resolution, it allows us also to manipulate the transport coefficients of the coarser model [128,156]. Formally, to derive the equations of motion for the coarse-grained DOFs, guaranteeing the dynamics defined by the fine-grained model, we have to resort to the Mori–Zwanzig formalism [157,158]. This results in the generalized Langevin equation [159]. Unfortunately, the latter is usually numerically unsolvable and we have to make several

approximations [160]. For instance, neglecting memory effects [160–163], we can employ the linear momentum preserving dissipative particle dynamics (DPD) thermostat [156,164] to match the transport coefficients, e.g., viscosity and diffusion constant, of fine- and coarse-grained models. This is, for instance, important in Open Boundary Molecular Dynamics (OBMD) simulations [81,82], presented in Section 3.2.2. In short, the Mori–Zwanzig formalism under Markovian approximation yields dynamics with DPD-like equations of motion for coarse-grained DOFs.

The DPD thermostat is introduced through the force $\mathbf{F}_\alpha^{thermo}$ as [156,164]:

$$\mathbf{F}_\alpha^{thermo} = \mathbf{F}_\alpha^{\mathrm{D}} + \mathbf{F}_\alpha^{\mathrm{R}}, \tag{53}$$

where $\mathbf{F}^{\mathrm{D}}$, and $\mathbf{F}^{\mathrm{R}}$ are damping and random forces, respectively [165,166]. The damping force is defined as

$$\mathbf{F}_{\alpha\beta}^{\mathrm{D}} = -\zeta\, w^{\mathrm{D}}(R_{\alpha\beta})\overleftrightarrow{P}_{\alpha\beta}(\mathbf{R}_{\alpha\beta})\mathbf{v}_{\alpha\beta}, \tag{54}$$

where $\mathbf{v}_{\alpha\beta} = \mathbf{v}_\alpha - \mathbf{v}_\beta$, and the random force as

$$\mathbf{F}_{\alpha\beta}^{\mathrm{R}} = \sigma w^{\mathrm{R}}(R_{\alpha\beta})\overleftrightarrow{P}_{\alpha\beta}(\mathbf{R}_{\alpha\beta})\boldsymbol{\theta}_{\alpha\beta}. \tag{55}$$

Here ζ and σ are the friction constant and the noise strength, respectively. The projection operator $\overleftrightarrow{P}$ is symmetric in the molecule indices ($\overleftrightarrow{P}_{\alpha\beta} = \overleftrightarrow{P}_{\beta\alpha}$). Contrary, the noise vector $\boldsymbol{\theta}_{\alpha\beta}$

$$\langle \boldsymbol{\theta}_{\alpha\beta}(t) \otimes \boldsymbol{\theta}_{\gamma\kappa}(t')\rangle = 2\overleftrightarrow{I}(\delta_{\alpha\gamma}\delta_{\beta\kappa} - \delta_{\alpha\kappa}\delta_{\beta\gamma})\delta(t-t'), \tag{56}$$

is antisymmetric ($\boldsymbol{\theta}_{\alpha\beta} = -\boldsymbol{\theta}_{\beta\alpha}$) in the molecule indices according to the fluctuation–dissipation theorem. The projection along the interatomic axis between molecule α and β, $\overleftrightarrow{P}_{\alpha\beta}(\mathbf{R}_{\alpha\beta}) = \hat{R}_{\alpha\beta} \otimes \hat{R}_{\alpha\beta}$, where $\hat{R}_{\alpha\beta}$ denotes a unit vector of the intermolecular axis $\mathbf{R}_{\alpha\beta} = \mathbf{R}_\alpha - \mathbf{R}_\beta$, then yields the standard DPD thermostat [164]. On the other hand, $\overleftrightarrow{P}_{\alpha\beta}(\mathbf{R}_{\alpha\beta}) = \overleftrightarrow{I} - \hat{R}_{\alpha\beta} \otimes \hat{R}_{\alpha\beta}$ gives the Transverse DPD thermostat [156].

Finally, the dynamics for a CoM velocity $\mathbf{V}_\alpha$ and coordinate $\mathbf{R}_\alpha$ of a given molecule α in AdResS MD simulations is governed by

$$\frac{d\mathbf{R}_\alpha}{dt} = \mathbf{V}_\alpha, \tag{57}$$

$$M_\alpha \frac{d\mathbf{V}_\alpha}{dt} = \mathbf{F}_\alpha^{AdResS} + \mathbf{F}_\alpha^{thermo}, \tag{58}$$

where M_α is the mass of the molecule α. Numerically, these equations of motion are then solved as explained in Appendix A.

At present, the AdResS scheme is implemented in two open source molecular dynamics packages: Extensible Simulation Package for Research on Soft matter systems (ESPResSo++) [167] and the GROningen MAchine for Chemical Simulations (GROMACS) [168].

3.1.3. Hamiltonian AdResS (H-AdResS) and related particle-based multiscale methods

AdResS, presented in the previous section, is constructed to satisfy Newton's Third Law and hence exactly preserves linear momentum. At the same time it provides the correct statistical and dynamical properties in both the atomistic and (trivially) the coarse-grained region. It must be noticed that the transition region is by definition an artificial domain whose only task is that of allowing a smooth transformation of molecules from a higher to a lower resolution and vice versa. Since the total AdResS force, given by Eq. (49), is non-conservative this prevents to define the global Hamiltonian of the system. Furthermore, as the number of DOFs changes on-the-fly, the method has to be used together with the local thermostat. On the other hand, a global Hamiltonian scheme would offer some technical benefits such as performing adaptive resolution MC simulations. However, from a conceptual point of view, if not properly used, this could lead to a misunderstanding of the meaning of physical properties considered (see discussion in Section 3.2.1). There have been some attempts in literature to devise an energy conserving adaptive resolution scheme [169–172]. As already discussed, this comes about with some methodological and conceptual difficulties [99–101]. Nevertheless, a method that comes closest to reach this goal is a variation of AdResS, i.e., H-AdResS [103,104,131,173–175].

Let us begin a brief presentation of H-AdResS by defining the total potential energy function, in analogy with Eqs. (38) and (40), as [103]

$$U = U^{int} + \sum_\alpha \left[\lambda_\alpha U_\alpha^B + (1-\lambda_\alpha)U_\alpha^A\right], \tag{59}$$

where

$$U_\alpha^A = \frac{1}{2}\sum_{\beta\neq\alpha} U^A(|\mathbf{R}_{\alpha\beta}|), \tag{60}$$

$$U_\alpha^B = \frac{1}{2}\sum_{\beta\neq\alpha}\sum_{i\alpha,j\beta} U^B(|\mathbf{r}_{i\alpha j\beta}|). \tag{61}$$

Here, U^{int} is the intramolecular potential energy, which is not subject to the interpolation, and U^A_α and U^B_α are the sums of all non-bonded intermolecular potentials corresponding to the low-resolution A and high-resolution B models, respectively, acting on a given molecule α. Note that now the coupling parameter $\lambda_\alpha = \lambda(R_\alpha)$ has only one subscript α and does depend only on the CoM of the molecule α and not on pairs of molecules, as in AdResS. We can hence write $\lambda(r) = w(r)$, i.e., the coupling parameter λ has the same functional form as the AdResS weighting function w given by Eq. (39). Such choice of coupling parameter is a consequence of the fact that in H-AdResS fine-grained DOFs are retained everywhere (also in the coarse-grained domain A, so that coarse-grained molecules close to the transition region interact also at the fine-grained level, which is not the case in AdResS with $\lambda_{\alpha\beta} = w_\alpha w_\beta$), i.e., the number of DOFs in the system is kept constant and only intermolecular interactions are changed while crossing the resolution domains. This is a necessary ingredient to keep the scheme time-reversible because molecules keep their orientations when going back and forth from the domains A and B [91]. (Moreover, it avoids the use of local thermostat to furnish or remove the latent heat associated with the turned off/on DOFs in the transition region between domains A and B [85].) This further implies that the form of kinetic energy is the same everywhere and we can therefore write the total Hamiltonian as [103]

$$H = T + U = \sum_\alpha T_\alpha + U^{int} + \sum_\alpha \left[\lambda_\alpha U^B_\alpha + (1-\lambda_\alpha)U^A_\alpha\right], \tag{62}$$

where $T_\alpha = \sum_{i\alpha} \frac{p^2_{i\alpha}}{2m_{i\alpha}}$ is the total kinetic energy of molecule α, summed up over all its atoms (combined subscript $i\alpha$ denotes ith atom of molecule α).

The corresponding force derived from Eq. (62) for ith atom of molecule α is [103]

$$\begin{aligned}\mathbf{F}_{i\alpha} &= \mathbf{F}^{int}_{i\alpha} + \sum_{\beta,\beta\neq\alpha}\left[\frac{\lambda_\alpha+\lambda_\beta}{2}\mathbf{F}^B_{i\alpha\beta} + \left(1-\frac{\lambda_\alpha+\lambda_\beta}{2}\mathbf{F}^A_{i\alpha\beta}\right)\right] + \frac{\partial\lambda_\alpha}{\partial\mathbf{r}_{i\alpha}}[U^A_\alpha - U^B_\alpha] \\ &= \mathbf{F}^{int}_{i\alpha} + \sum_{\beta,\beta\neq\alpha}\left[\frac{\lambda_\alpha+\lambda_\beta}{2}\mathbf{F}^B_{i\alpha\beta} + \left(1-\frac{\lambda_\alpha+\lambda_\beta}{2}\mathbf{F}^A_{i\alpha\beta}\right)\right] + \mathbf{F}^{dr}_{i\alpha},\end{aligned} \tag{63}$$

where $\mathbf{F}^{int}_{i\alpha}$ is the force acting on atom $i\alpha$ due to other atoms in the same molecule, $\mathbf{F}^A_{i\alpha\beta}$ and $\mathbf{F}^B_{i\alpha\beta}$ are coarse-grained and fine-grained forces, respectively, acting on atom $i\alpha$ due to interaction with molecule β, and $\mathbf{F}^{dr}_{i\alpha}$ is the drift force. The above force given by Eq. (63) is, not surprisingly, very similar to the force given by Eq. (41) and $\mathbf{F}^{dr}_{i\alpha}$ to the drift force given by Eq. (44). As explained in the previous subsection about AdResS, the drift force $\mathbf{F}^{dr}_{i\alpha}$ violates Newton's Third Law and hence so does the total force $\mathbf{F}_{i\alpha}$. Because the drift forces pushes molecules in the domain with a lower chemical potential, the density profile obtained by the energy-conserving and linear-momentum non-conserving dynamics governed by Eqs. (62) and (63) is not uniform across the simulation box. One encounters similar density fluctuations [149] in the transition region as in AdResS, Eq. (45), without application of the thermodynamics force. Therefore, the microscopic global mixed resolution Hamiltonian, Eq. (62), does not describe a homogeneous system with uniform density profile that one would like to model.

To enforce a uniform density across the system, a new term is introduced in the mixed-resolution Hamiltonian above, Eq. (62), resulting in a modified global Hamiltonian [103]

$$\hat{H} = H - \sum_\alpha \Delta H(\lambda_\alpha). \tag{64}$$

The additional term changes the drift force to [103]

$$\hat{\mathbf{F}}^{dr}_\alpha = \frac{\partial\lambda_\alpha}{\partial\mathbf{R}_\alpha}\left[U^A_\alpha - U^B_\alpha + \frac{d\Delta H(\lambda)}{d\lambda}|_{\lambda=\lambda_\alpha}\right]. \tag{65}$$

To remove, on average, the drift force, i.e., $\langle\hat{\mathbf{F}}^{dr}\rangle = 0$, using Eq. (36), the additional term must satisfy [103]

$$\frac{d\Delta H(\lambda)}{d\lambda}|_{\lambda=\lambda_\alpha} = -\langle U^A_\alpha - U^B_\alpha\rangle_{\mathbf{R}_\alpha}. \tag{66}$$

This leads to the uniform hydrostatic pressure profile ($\nabla p = \rho\langle\hat{\mathbf{F}}^{dr}\rangle = 0$) across the simulation box while in domains A and B the density still may differ due to different equations of state of both models [103]. Thus, to get the uniform density profile across the simulation box [174], eventually, one has to resort to the same old thermodynamic force $\mathbf{F}^{TD}_\alpha = \mathbf{F}^{TD}(R_\alpha)$ obtained in the iterative way, as explained in the previous subsection, by Eq. (52).

The last remark should not come as a surprise, as the "thermodynamic" correction term $\sum_\alpha \Delta H(\lambda_\alpha)$ is state point dependent and does not directly couple to the microscopic DOFs [175]. Hence, the modified global Hamiltonian $\hat{H}$ is not microscopic and furthermore, it is state dependent, as it carries an *a priori* calculated free energy in the correction term. Following basic principles of statistical mechanics, free energy should be a quantity calculated from the Hamiltonian of the system, which expresses the physics of the system, and not the other way round. In any case, provided that this does not present any conceptual issues, the derived H-AdResS, at least in technical terms, opens doors to the technical possibility of equilibrium simulations in different statistical mechanics ensembles [131] as well as to adaptive resolution

MC techniques [104]. The method is energy ($\hat{H}$) conserving, in the sense that the defined energy remains constant during the simulation, and despite broken translational invariance and associated linear momentum non-conservation it has been shown that the method is thermodynamically translational invariant [132,175].

To conclude this discussion of adaptive resolution schemes, let us report, at this point, some related concurrent particle-based multiscale approaches that have been developed in the past few years. Abrams carried out the first concurrent dual-resolution MC simulation [169]. The method was applied to a system of liquid methane at a single thermodynamics state point, where half of simulation box was filled with atomistic molecules of liquid methane whereas the other half was occupied with corresponding spherically symmetric united atoms, interacting via an effective coarse-grained pairwise intermolecular potential. Christen et al. developed Multigraining, a method to combine fine- and coarse grained resolutions. Similarly to AdResS or H-AdResS, a coupling parameter λ is introduced to couple the fine- and coarse-grained potentials in the total potential energy [176]. However, the crucial difference between the methods is that λ in Multigraining is not position-dependent, as in AdResS or H-AdResS, preventing the possibility of performing adaptive resolution simulations. Instead, Multigraining is an efficient method for fast equilibration of molecular liquids, e.g., liquid octane a hexadecane [176], or replica-exchange simulations, where λ is used to distinguish between the different replicas. Ensing et al. presented a concurrent multiscale MD algorithm, representing an attempt to devise an energy-conserving adaptive resolution MD method [171]. To this end, they introduced a "bookkeeping" energy, whose corresponding force should exactly cancel the drift introduced by the interpolated potential, similarly to Eq. (65) in H-AdResS. This leads to a conceptual contradiction, as the resulting total force has a non-conservative AdResS-like form [101]. Nevertheless, combining this method with a rotational dynamics of rigid atomistic fragments in the low resolution domain allowed for the first MD simulation of a solvated macromolecule (a polyethylene chain in liquid hexane), where different parts of the macromolecule could be represented at the same time with different levels of details, changing resolution adaptively, on-the-fly [172]. This is different to some other approaches that combine atomistic and coarse-grained resolutions within a macromolecule, e.g., a synthetic polymer next to a metallic surface [177] or a globular protein [178]. There, some relevant region of the macromolecule is treated with the atomistic level of detail whereas the rest of the macromolecule is represented with the coarse-grained model. However, the atomistically represented part of the macromolecule remains atomistic during the entire course of simulation. The same holds also for the coarse-grained part. A very interesting adaptive resolution simulation approach was introduced by Heyden et al. [170,179,180]. The system is partitioned into an active zone, i.e., the region of interest, the environmental zone (low resolution domain), and a buffer zone, separating the active from the environmental zone. In the first version of the method, called Adaptive partitioning (AP) method [170], the potential energy function is defined as a linear combination of all possible combinations of multiresolution energies that are obtained by treating the region of interest and a subset of groups of atoms in the buffer zone at a high level of theory (e.g., quantum mechanical) and the rest at a low level of theory (e.g., molecular mechanical). The method was applied to simulation of solvated Li^+ ion and a supercritical fluid of argon atoms. The improved version, called Adaptive partitioning of the Lagrangian (APL) method [179], extends the AP method to Lagrangian functions, i.e., the mixed-resolution Lagrangian is defined as a linear combination of all possible combinations of Lagrangian functions that are obtained by treating the active zone and a subset of atom groups in the buffer zone at a high level of resolution and the remainder at a low level of resolution. This further allows one to define a mixed-resolution Hamiltonian. Even though the APL method certainly represents an innovative and promising approach to adaptive resolution simulation, it poses a major challenge for algorithmic implementation in MD or MC computer packages. Hence, its only application to date is to a system of supercritical methane [180]. A promising approach based on extended Lagrangian formalism that allows for a smooth transition in the time domain between two different interaction potentials during an MD simulation has been recently introduced in Ref. [181]. For example, a classical force field (MM) MD simulation can be dynamically switched on demand into a hybrid QM/MM-like simulation once one changes the MM treatment of the particular subdomain of the system into a QM subsystem and vice versa. Another group of methods couples atomistic with coarse-grained force fields keeping the resolution of each particle fixed during the simulation. Some of these fixed resolution approaches make use of virtual sites [182–185]. The virtual sites are defined for groups of fine-grained atoms based on particular coarse-grained mapping, e.g., MARTINI force field. Thus, the coarse-grained molecules only interact with these virtual sites and not with the fine-grained atoms. Then, using the virtual sites one defines the mixed interactions between atomistic and coarse-grained molecules in a straightforward way, such that they are treated either the same way as the pure intermolecular interactions between coarse-grained molecules or they are slightly reparametrized. The interactions between molecules of the same resolution are described by respective pure force-fields without further alternations. Similar fixed resolution approaches are reported in Refs. [186–190]. This kind of methods has been applied to variety of biomolecular systems. Another interesting approach, named Relative Resolution (RelRes), has been introduced in Ref. [191]. In RelRes, molecules interact with each other via fine-grained interactions at small relative separations whereas coarse-grained potentials are used for intermolecular interactions at large relative separations. This approach can be understood as an extension of force truncation [191]. Riniker et al. [192] mixed atomistic and supramolecular coarse-grained water in the MD simulation of liquid water. They parameterized the mixed fine- and coarse-grained interactions to reproduce thermodynamic and dielectric properties of liquid water at ambient conditions for various mole fractions of coarse-grained in atomistic water. Finally, let us also mention adaptively restrained particle simulations (ARPS), introduced by Artemova et al. [193]. ARPS allows us to speed up particle simulations by adaptively switching on and off positional DOFs, while allowing momenta to evolve. This is achieved by introduction of a Hamiltonian with a modified inverse inertia matrix, which is used to specify how and when the positional DOFs are switched on/off during the simulation.

3.2. Open boundary molecular simulations

As already discussed in the Introduction, many physical systems are open, i.e., mass, momentum, and energy are exchanged with the surroundings. However, standard MD simulations as well as adaptive resolution simulation approaches presented in the previous subsection are typically performed using periodic boundary conditions with a constant number of particles. In other words, the studied systems are closed without the possibility of using arbitrary external boundary conditions, which would allow us to incorporate into our model the transfer of mass, momentum, and energy across the boundaries of the studied molecular system. In the following, we present a few simulation approaches, based on AdResS, to go beyond this limitation and make it possible the molecular simulations of open systems, either in thermodynamic equilibrium in the GC ensemble or under arbitrary non-equilibrium conditions.

3.2.1. AdResS within the framework of the Grand Canonical ensemble

The main criticism to the standard AdResS method is that, because it is based on interpolation of forces, it does not allow to define a global Hamiltonian. Hence, the averages done in the high resolution region, while numerically sound, should not be justified from a statistical mechanics point of view [103,171,179]. It must be clarified first, that a global Hamiltonian in an adaptive scheme (such as that of Section 3.1.3), while being technically useful for plugging in the standard algorithms of MD and MC in a straightforward way, leads to a conceptual misunderstanding if employed as a justification for definition of a global canonical ensemble [131,194,195]. The basic principle of an adaptive resolution simulation states that the aim of such simulations is to have, in the high-resolution region of the adaptive setup, the same properties as a full highly-resolved system. This implies that average properties done in the high resolution region should be computed over a subsystem whose number of molecules is variable. Moreover, since the atomistic (or high resolution) DOFs in the low-resolution remainder of the adaptive setup are either completely integrated out (as in AdResS) or treated only to define the intramolecular energy (as in H-AdResS), the statistical average of a quantity (where the atomistic DOFs are required) computed globally (and thus requiring a global Hamiltonian that defines a global canonical ensemble) may not necessarily have a physical meaning by definition. For example, let us consider a system of liquid water (using a flexible atomistic water model) in the H-AdResS setup, where atomistic intramolecular DOFs are defined everywhere (also in the low resolution domain). This means that one can define dipole moments of individual water molecules throughout the entire simulation box. Now, using a global Hamiltonian defined by Eq. (64) and computing a statistical average value of the dipole moment (in the corresponding global canonical ensemble) would yield some average value. However, this average would not match the corresponding average value from the full-blown atomistic simulation, as the orientations of water molecules in the low resolution domain are false due to lack of atomistic (electrostatic) intermolecular interactions. Thus, the search for a statistical mechanics justification of the properties, calculated in the high-resolution subdomain, should follow a different path than that of a global Hamiltonian.

Indeed, the search for formal justification in terms of statistical ensemble for quantities, calculated in the high-resolution (atomistic) region, has been subject of intensive investigations in the last years. The definition of the thermodynamic force, on the basis of the empirical principle that different regions at different resolution should have the same chemical potential (in order to be in thermodynamic equilibrium), represented the first step towards a more rigorous definition of the thermodynamics and statistical mechanics of the atomistic region. An important step forward was done in Ref. [133]. There, the thermodynamic force is derived following the principle that two open systems, exchanging energy and matter, to be in equilibrium must satisfy (at equal volume) the equivalence of the grand potential. Moreover, the transition/hybrid region can be considered as a technical/artificial thin filter, which does not perturb the physics of the atomistic region. Furthermore, it allows for the proper exchange of energy and matter and for transformation of molecules from a coarse-grained reservoir. The formal expression results in: $p_{AT} + \rho_0 \int_{\Delta} \mathbf{F}^{TD}(\mathbf{r}) d\mathbf{r} = p_{CG}$, where p_{AT} is the chosen pressure of the atomistic system (region), p_{CG} is the pressure of the coarse-grained model, ρ_0 is the chosen molecular density of the atomistic system (region). From this point on, the question is, whether the idea of balance of grand potential and having the coarse-grained plus hybrid region as a reservoir of energy and matter can be formalized more precisely in terms of the atomistic region as a GC ensemble. In fact, if the atomistic region can be considered in the GC ensemble, then all the averages done in such a region are rigorously defined in terms of statistical mechanics, as clearly underlined also in Ref. [175].

This idea was explored in Refs. [196,197]. It was found that the thermodynamic force in terms of gradient of the number density ρ expresses a mathematically necessary and numerically sufficient condition to have a GC distribution in the atomistic region, independently of the coarse-grained model. The latter can be any liquid of spherical particles with the only (technical) constrain that it is at the same density as the liquid in the atomistic region. Moreover, the sum of the work of the thermodynamic force in the transition region, plus the energy provided by the thermostat, $\omega^{TD} + \omega_Q$ (corresponding to the latent heat of the previous sections) equals the difference of chemical potential between the atomistic and coarse-grained regions, $\mu_{CG} - \mu_{at}$. A quantitative proof, that the concept introduced above was very solid, was done by calculating the excess chemical potential of various liquids and mixtures, following the procedure outlined above [198]. Table 1, reports the satisfactory results found. Due to the fact that the coarse-grained region could be represented by any generic liquid of spheres and that the attention of our physical analysis was focused on the atomistic region only (as a GC-like ensemble), we renamed this approach "GC-AdResS". However, even in the detailed mathematical analysis of Ref. [197], the definition of Hamiltonian of interest is not straightforward. To define the Hamiltonian of the atomistic region, one has to use the following technical operation: part of the atomistic region is included in the transition region, so that the interaction between a smaller atomistic region (now the region of interest) and the rest of the system is restricted to a well defined standard atomistic interaction only

Table 1
The excess chemical potential for different mole fractions of solute (x) calculated following the GC-AdResS procedure and the thermodynamic integration (TI) technique of full atomistic simulations. The comparison with the standard technique used in MD (TI) and (when available) with experiments shows a very satisfactory agreement.
Source: Reprinted from Ref. [198].

Liquid component	x	GC-AdResS	TI	Experiment
Water	–	-22.8 ± 0.2	-22.1 ± 0.3	-23.5
Methane	–	-4.6 ± 0.1	-5.2 ± 0.1	–
Ethane	–	-8.2 ± 0.3	-8.8 ± 0.1	–
Propane	–	-8.5 ± 0.1	-9.5 ± 0.2	–
Methanol	–	-20.1 ± 0.1	-20.6 ± 0.4	-20.5
DMSO	–	-32.2 ± 0.3	-34.7 ± 0.7	-32.2
Methanol in methanol/water mixture	0.01	-18.1 ± 0.2	-19.7 ± 0.2	–
Methane in methane/water mixture	0.006	9.1 ± 0.1	8.5 ± 0.2	–
Urea in urea/water mixture	0.02	-56.1 ± 0.6	-58.2 ± 0.5	-57.8 ± 2.5
Ethane in ethane/water mixture	0.006	7.2 ± 0.2	7.4 ± 0.3	–
TBA in water/TBA mixture	0.001	-19.5 ± 0.3	-20.8 ± 0.6	-19.0
DMSO in DMSO/water mixture	0.01	-31.4 ± 0.5	-33.2 ± 0.3	–
TBA in TBA/DMSO mixture	0.02	-24.8 ± 0.4	-24.0 ± 0.5	–

(due to the cutoff of the atomistic interaction). Because of this, it can be written analytically. This procedure is conceptually acceptable and numerically valid but not optimal, because it implies the loss of some computational gain. Moreover, it does not provide a direct check of the main hypothesis, on which the GC ensemble is built in statistical mechanics, i.e., the surface energy must be negligible w.r.t. the volume energy of the atomistic region (see e.g. the derivation of K. Huang [58]). This aspect was treated in a following work [199] and a solution was found by explicitly identifying the interaction energy of the molecules of the atomistic region with the reservoir (transition region). In the following, we will consider the atomistic region as the system of interest and the rest as reservoir. To this end, we identify the models A and B of Section 3.1.2 with the coarse-grained (CG) and atomistic (AT) regions, respectively, while the transition region is identified with the hybrid (HY) region, the CG+HY region is defined as the reservoir (RES). For molecule α, at position $\mathbf{R}_\alpha$ in the AT region of AdResS, one has $w_\alpha = w(\mathbf{R}_\alpha) = 1$. The corresponding force can be partitioned in two parts, i.e., (a) the force generated by the interaction of molecule α with molecules of the AT region:

$$\mathbf{F}_{\alpha\beta} = \mathbf{F}^{AT}_{\alpha\beta}, \forall \beta \in \quad AT \tag{67}$$

and (b) the force generated by the interaction with the molecules in the reservoir:

$$\mathbf{F}_{\alpha\beta} = w_\beta \mathbf{F}^{AT}_{\alpha\beta} + [1 - w_\beta]\mathbf{F}^{CG}_{\alpha\beta}, \forall \beta \in \quad HY + CG. \tag{68}$$

From Eq. (67) one can express the force acting on molecule α in terms of the gradient of the atomistic potential:

$$\mathbf{F}_\alpha = \sum_{\beta \neq \alpha} \mathbf{F}^{AT}_{\alpha\beta} = -\sum_{\beta \neq \alpha} \nabla_\alpha U^{AT}, \tag{69}$$

where ∇_α is the gradient w.r.t. molecule α. Eq. (68) corresponds to the action of the molecules in the reservoir on molecule α, that is an external force (w.r.t. the AT region). Eq. (68) can be rewritten as:

$$\mathbf{F}_\alpha = \sum_{\beta \in HY+CG} [w_\beta \mathbf{F}^{AT}_{\alpha\beta} + [1 - w_\beta]\mathbf{F}^{CG}_{\alpha\beta}] = -\sum_{\beta \in HY+CG} [w_\beta \nabla_\alpha U^{AT} + [1 - w_\beta]\nabla_\alpha U^{CG}]. \tag{70}$$

Eq. (70) implies that the net force on molecule α can be considered as a (non-local) gradient field instantaneously generated by the external field due the other molecules (of the reservoir). It follows that the energy of molecule α at time $t > 0$ associated with the coupling force can be defined as:

$$W^\alpha_{AT-RES}(t) = \sum_{\beta \in HY+CG} [w_\beta U^{AT}_{\alpha\beta} + [1 - w_\beta]U^{CG}_{\alpha\beta}], \tag{71}$$

where the $U_{\alpha\beta}$ is the interaction energy between molecule α at position $\mathbf{R}_\alpha$ and the other molecules sitting at a certain $\mathbf{R}_\beta$. The total $AT - RES$ energy of the molecules in the atomistic region at time t is then defined as:

$$W_{AT-RES}(t) = \sum_{\alpha \in AT} W^\alpha_{AT-RES}(t)\,. \tag{72}$$

The quantity of Eq. (72) is nothing else than the surface energy of [58] and must be compared to the amount of energy, W_{AT-AT}, corresponding to the interaction between molecules of the AT region only (volume energy in Ref. [58]): $W_{AT-AT}(t) = \sum_{\alpha<\beta} U^{AT}_{\alpha\beta}$; $\alpha, \beta \in AT$. From the formal and computational point of view, if:

$$\frac{|W_{AT-AT}(t)| - |W_{AT-RES}(t)|}{|W_{AT-AT}(t)|} \approx 1; \forall t \tag{73}$$

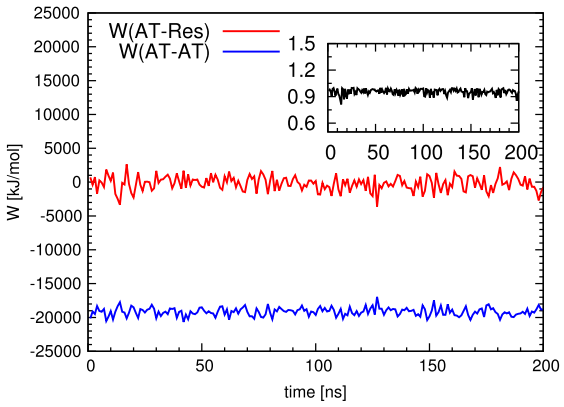

Fig. 5. Main figure: Interaction energy of the atomistic region only as a function of time, $W_{AT-AT}(t)$ compared to the interaction energy between the atomistic region and reservoir, $W_{AT-RES}(t)$. The energy of the atomistic region is at least one order of magnitude than the atomistic-reservoir energy. Inset: The ratio $\frac{|W_{AT-AT}(t)|-|W_{AT-RES}(t)|}{|W_{AT-AT}(t)|}$ as a function of time. It is evident that the surface effect is, at most, of 10%. However, in other tests done with larger systems of typical interest in AdResS, the effect goes even below 1.0%.
Source: Reprinted from Ref. [199].

then the surface effects are negligible w.r.t. the volume effects and thus it seems reasonable to approximate the total energy of the atomistic region by the Hamiltonian of the AT region only,

$$H_{AT} \approx H_{AT-AT}. \tag{74}$$

Fig. 5 shows that, for liquid water, even for relatively small atomistic regions, sizable hybrid regions and relatively small coarse-grained region (i.e. a worst-case-scenario from the technical point of view) condition (73) holds.

In general, the importance of condition (73) is that it provides a practical criterion to check whether or not one is in the conditions for treating the atomistic region in a GC fashion, i.e., if the method, in this sense, is truly rigorous. This shows that a global Hamiltonian is not needed in order to justify the statistical averages done in the atomistic region. Moreover, the design of the reservoir can be arbitrarily done and simplified to a very basic level, as long as the conditions of the BL model are satisfied. For example, in the particular case of numerical studies of liquid water reported here, the coarse-grained region is composed of a liquid of simple soft spheres. Thus, it does not have any structural similarities (i.e. radial distribution function) with the actual liquid of the atomistic region. The possibility of such simplifications paves the way for a simple and efficient design of an interface with, e.g., hydrodynamic approaches (see Section 3.2.2), and thus, as a matter of fact, GC studies with infinite reservoirs.

Finally, as reported in Section 2.1 and in Refs. [199,46], the calculation of equilibrium time correlation functions in the GC ensemble requires the formalism of Bergman and Lebowitz. In Ref. [199], it has been shown a plausible correspondence between the BL model and GC-AdResS. A fully rigorous procedure would require the writing of the BL kernel $K_{NM}(X'_N, X_M)$, reported in Section 2.1, explicitly in terms of the forces of AdResS. However, this would be a formidable task at this stage. Instead, in a simpler, yet satisfactory way for the justification of the numerical approach, the correspondence was done by (a) identifying the Hamiltonian of the BL model with the Hamiltonian of the atomistic region, H_{AT-AT}, (b) by then showing that the thermodynamic force in GC-AdResS enforces the BL condition $\sum_{N=0}^{\infty} \int dX'_N [K_{MN}(X_M, X'_N)\rho(X'_N, N, t) - K_{NM}(X'_N, X_M)\rho(X_M, M, t)] = 0$, and finally (c) the definition of the chemical potential required by the BL model is achieved through the work of the thermodynamic force plus that of the thermostat. Furthermore, we checked that the conditions of thermodynamic invariance of the reservoir in time and of symmetric exchange of particles required by the BL model were satisfied in GC-AdResS. It is reasonable then to consider the BL model as the conceptual mathematical and physical basis for GC-AdResS so that it can be considered an open boundary GC simulation model. The consequence is the operational definition of time correlation function of the AT region in GC-AdResS, as reported in Section 2.4. Following such a definition, numerical tests were then carried on for various time correlation functions and the results were rather satisfactory (see, e.g., Fig. 6). It must be reported that, to avoid artificial effects in the dynamics of the atomistic region, the thermostat acts only in the reservoir (i.e. transition region plus the coarse-grained region). The question remaining for the readers of this work is, where to go from here. The task of writing, in a rigorous way, the BL kernel explicitly, in terms of MD computable quantities, would be a rather desirable result. At this stage, we can only speculate that if a solution to this problem is found, any open boundary approach, based on MD, would necessarily converge to such a solution leading to a general algorithm of universal character. The work reported here, can be envisaged only as a starting point, although already sufficient for numerical applications.

3.2.2. Open Boundary Molecular Dynamics (OBMD)

An open boundary methodology that truly opens up the simulation domain and allows for molecules exchange is the OBMD method [81,82], which is a combination of two methods, i.e., open MD [200,201] and AdResS. Open MD enables us

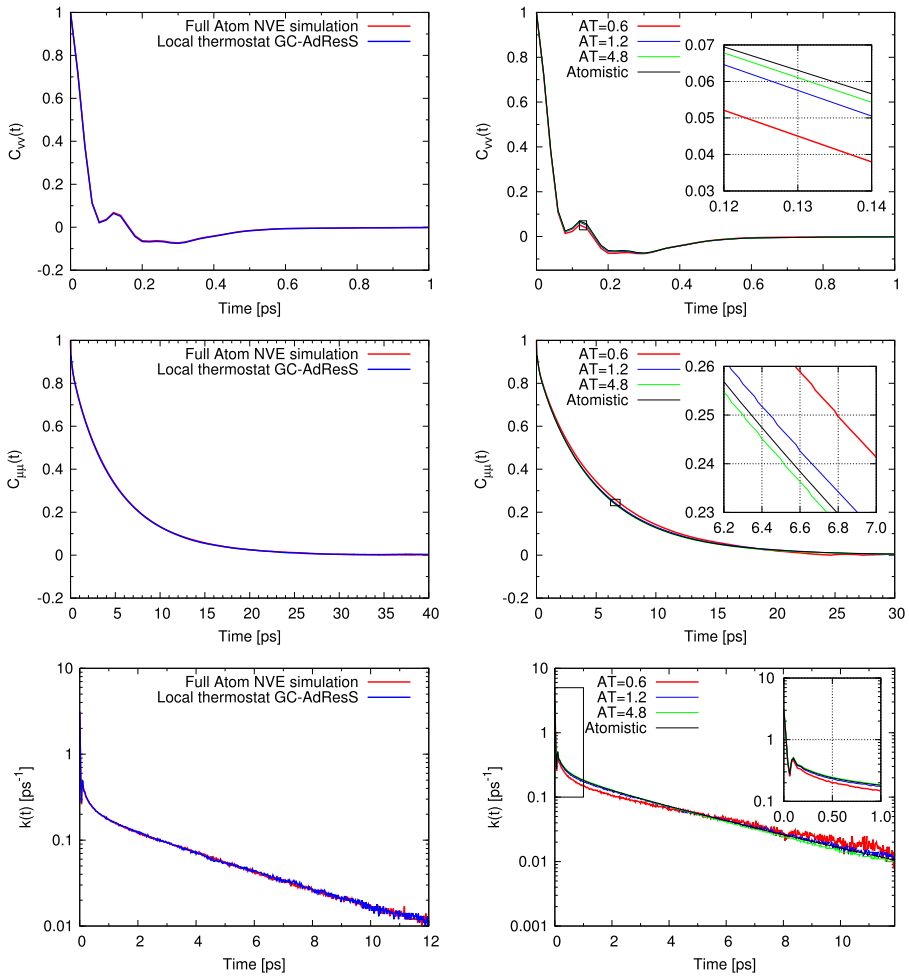

Fig. 6. (Left column) Three equilibrium time correlation functions for liquid water using the SPC/E model at room conditions. The functions are calculated using GC-AdResS and, in an equivalent subsystem, a fully atomistic simulation. The functions are defined as: velocity–velocity autocorrelation function, $C_{VV}(t)$, (molecular) dipole–dipole autocorrelation function, $C_{\mu\mu}(t)$, reactive flux correlation function, $k(t)$ (semilogarithmic plot). The agreement between GC-AdResS and the fully atomistic simulation is highly satisfactory. Reprinted from Ref. [199]. (Right column) Systematic convergence of $C_{VV}(t)$, $C_{\mu\mu}(t)$ and $k(t)$ (semilogarithmic plot) calculated using GC-AdResS to the fully atomistic results calculated over the whole system (NVE simulation). Here, the principle of equivalence of ensembles in the thermodynamic limit (GC and microcanonical in this case) is checked by systematically increasing the AT region in a large GC-AdResS simulation. The obtained results are compared with a full atomistic simulation of a large system in the microcanonical (NVE) ensemble. This result, together with that of the previous figure, shows the consistency of the operational definition of time correlation function given by the BL model within the GC-AdResS numerical setup.
Source: Reprinted from Ref. [199].

to insert heat and momentum through the boundary of computational domain by external forces acting on the adjacent buffer domains. To allow the mass exchange buffers act also as molecules reservoirs, where molecules are deleted and inserted facilitated by AdResS, as explained shortly. A representative OBMD(≡ open MD+AdResS) system is depicted in Fig. 7, showing an OBMD setup of star-polymer melt. There, relatively large star-polymers molecules freely flow inwards or outwards of the simulation box according to the externally imposed thermo-mechanical state. The central MD domain (region of interest) is sandwiched between two buffer domains [82]. The latter two allow the former to exchange mass, momentum, and energy through two of its boundaries with the buffers along the horizontal x-direction. Thus, the OBMD system is opened, i.e., not periodic, in this direction. The star polymers freely move between the MD and buffer domains. Additionally, in the buffers, the resolution change from the fine- (next to the MD domain) to the coarse-grained resolution (at the outer boundaries of the simulation box) takes place. In the coarse-grained parts of the buffers, a given star polymer, consisting of 73 monomers, is represented with only one very soft coarse-grained bead. Here, we face the same problem as in computation of chemical potential [67,202], namely the difficulty of inserting the new molecules into a dense liquid. Thus, the idea behind the resolution change in the buffers is that AdResS allows the insertion of molecules of arbitrary size into the system. The coarse-grained domains of the buffers act as a mass reservoir where large molecules can be easily inserted due to soft effective interactions among coarse-grained beads [203–205] (see below). Then, as the molecules move toward

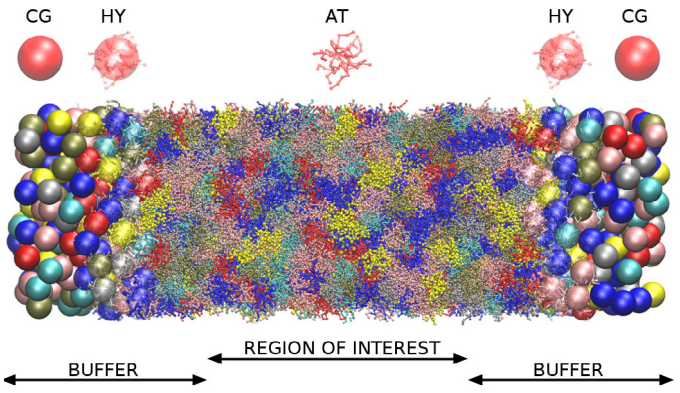

Fig. 7. OBMD setup of a star-polymer melt. The MD domain is embedded with two buffer domains, which serve as mass and momentum reservoirs (see text). For technical and other details the reader is referred to Refs. [81,82].
Source: Reprinted from Ref. [82].

the MD domain, they gain the fine-grained details employing AdResS in the buffers. Molecules are deleted once they leave the outer boundary of a given buffer and new molecules are inserted to achieve the mass balance, i.e., to have a desired average density in the buffer. The mass balance is controlled by a feedback algorithm, $\Delta N_B = (\Delta t/\tau_r)(\langle N_B\rangle - N_B)$, where $\langle N_B\rangle$ and N_B are the average and the current number of molecules in the buffer, while τ_r is the characteristic relaxation time of the buffer, typically of the order $\tau_r \sim O(100)$ MD timesteps. A molecule is deleted if $\Delta N_B < 0$ or when the molecule leaves the simulation box at the outer boundary of the buffer. New molecules are inserted if $\Delta N_B > 0$. The insertion at the coarse-grained domain of the buffers is carried out by an iterative algorithm USHER, which is a Newton–Raphson-like search method on the potential energy surface [203,204].

To impose boundary conditions to the computational domain, OBMD resorts to external force $\mathbf{F}^{ext}$ acting on a given buffer B with the interface of surface A to the MD domain. $\mathbf{F}^{ext}_{tot}$ is computed from the momentum and energy flux balance [200,206]

$$\mathbf{J}_p \cdot \mathbf{n} A \Delta t = \mathbf{F}^{ext}_{tot} \Delta t + \sum_{\alpha'} \Delta(M_{\alpha'} \mathbf{V}_{\alpha'}) \tag{75}$$

$$\mathbf{J}_e \cdot \mathbf{n} A \Delta t = \sum_{\alpha \in B} \mathbf{F}^{ext}_{\alpha} \cdot \mathbf{V}_{\alpha} \Delta t + \sum_{\alpha'} \Delta \epsilon_{\alpha'}, \tag{76}$$

where $\mathbf{J}_p$ and $\mathbf{J}_e$ are the momentum flux tensor and heat flux vector, respectively, that one would like to impose across each MD boundary over timestep Δt, and $\mathbf{F}^{ext}_{tot} = \sum_{\alpha \in B} \mathbf{F}^{ext}_{\alpha}$. Here, $\mathbf{n}$ is the unit vector normal to the buffer interface (along the x-direction in Fig. 7) and index α' runs over molecules that have entered or exited the buffer in the last timestep Δt. Hence, for the momentum change we have that $\Delta(M_{\alpha'} \mathbf{V}_{\alpha'}) = \pm M_{\alpha'} \mathbf{V}_{\alpha'}$ if the molecule α' enters (+) or leaves (−). The corresponding energy change is $\Delta \epsilon_{\alpha'} = \pm \epsilon_{\alpha'}$. Thus, these terms of Eqs. (75) and (76) measure the momentum and energy release due to molecule exchange with the surrounding. Eqs. (75) and (76) ensure that the total linear momentum and energy are conserved [81]. In the following, we will devote our attention on the linear momentum balance only. For further discussion on the energy transfer, please read Refs. [200,201].

Next, we have to redistribute the total external force $\mathbf{F}^{ext}_{tot} = A\left(\mathbf{J}_p \cdot \mathbf{n} - \frac{\sum_{\alpha'} \Delta(M_{\alpha'} \mathbf{V}_{\alpha'})}{A \Delta t}\right) = \sum_{\alpha \in B} \mathbf{F}^{ext}_{\alpha}$, which exactly conserves the linear momentum of the whole molecular system (buffers + MD domain), at each particular time among the molecules in the buffer B ($\alpha \in B$) with $\mathbf{F}^{ext}_{\alpha} = 0$ outside B ($\alpha \notin B$). Thus,

$$\mathbf{F}^{ext}_{\alpha} = \mathbf{G}(\mathbf{R}_{\alpha}) \mathbf{F}^{ext}_{tot}, \tag{77}$$

where $\mathbf{G}(\mathbf{R}_{\alpha})$ is a weighting function, i.e., $\sum_{\alpha \in B} \mathbf{G}(\mathbf{R}_{\alpha}) = \mathbf{1}$, with, in general, tensorial form (to distribute normal and tangential forces) [81].

Now, we are all set to write down equations of motion for a given molecule α in the OBMD system (MD domain+two buffers). Exploiting Eqs. (49), (53), (77) and (58) the dynamics is governed by

$$\frac{d\mathbf{R}_{\alpha}}{dt} = \mathbf{V}_{\alpha}, \tag{78}$$

$$M_{\alpha} \frac{d\mathbf{V}_{\alpha}}{dt} = \mathbf{F}^{AdResS}_{\alpha} + \mathbf{F}^{thermo}_{\alpha} + \mathbf{F}^{ext}_{\alpha}. \tag{79}$$

Since the change of molecular resolution takes place in the buffers and not in our region of interest, i.e., the MD domain, we set $\mathbf{F}^{TD}_{\alpha} = 0$ in Eq. (49). We are allowed to do this because the density fluctuations due to coupling of fine- and

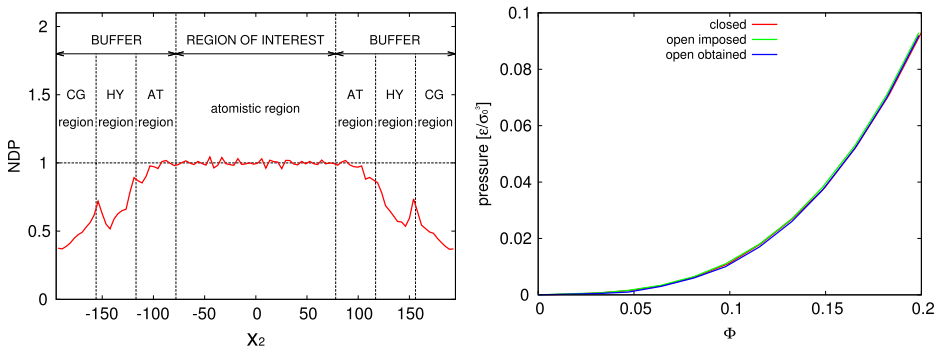

Fig. 8. (Left) Normalized density profile across the simulation box. In the MD domain (region of interest), the density profile is homogeneous and density reaches the bulk value. There is a noticeable density undulation present at the interface between the coarse-grained (CG) and transition (HY) regions in the buffer because there is no thermodynamic force applied (see text). (Right) Equation of state obtained from closed (with periodic boundary conditions in all three directions applied) and OBMD simulations in thermodynamic equilibrium. $p(\rho)$ curve in closed systems is obtained by varying system's density and computing the corresponding pressure from relaxed equilibrium state. In open system, $p(\rho)$ is calculated by changing externally imposed pressure and monitoring the density to which the system relaxes. All equations of state match well within the error bars.
Source: Adapted and reprinted from Refs. [81,82].

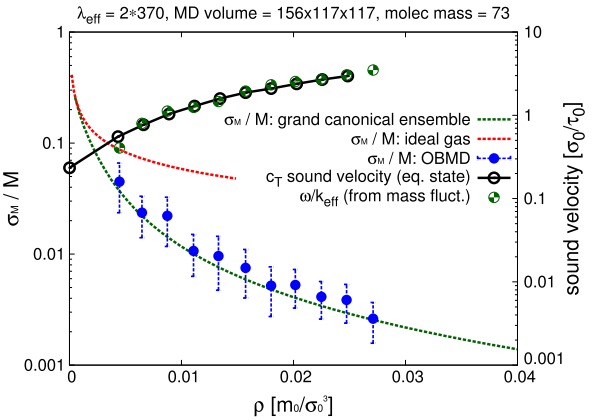

Fig. 9. Relative mass fluctuation of the MD domain of the OBMD setup from Fig. 7. GC theoretical prediction $\mathrm{Std}[M]/M = [k_BT/(Mc_T^2)]^{1/2}$ and the ideal gas limit using the isothermal sound velocity $c_T^{id} = (k_BT/M_m)^{1/2}$ are also shown. The right ordinate axis shows the values of the isothermal sound velocity $c_T = (\partial p/\partial\rho)_T$ using equation of state from Fig. 8 (right) and compared with ω/k_{eff} computed from the oscillation frequency ω of the total mass autocorrelation function in the MD domain.
Source: Reprinted from Ref. [82].

coarse-grained models using AdResS are small in comparison with the overall density change in the buffers, see Fig. 8 (left). Thus, $\mathbf{F}_\alpha^{AdResS}$ includes only pairwise intermolecular interactions as does the $\mathbf{F}_\alpha^{thermo}$ of the DPD thermostat, both preserving linear momentum. The nonphysical density wave reflections that might occur can be reduced employing non-reflecting boundaries [207].

As already stated, OBMD allows us to impose the external pressure tensor and the heat flux across the boundaries of the computational domain, see Fig. 8 (right). The external pressure and temperature determine the external chemical potential via the Gibbs–Duhem route. These are the independent parameters of the Grand-Canonical ensemble. Therefore, OBMD allows us to perform Grand-Canonical MD simulations with a fluctuating number of molecules. The fluctuations in the number of molecules inside the MD domain are consistent with the prediction from the GC ensemble, as shown in Fig. 9. Extended ensembles under non-equilibrium stationary states and time-dependent forcing can also be simulated using OBMD [81,82]. OBMD can be also used as an interface to connect MD with continuum hydrodynamics in hybrid particle-continuum methods [208]. This will be discussed in more detail in Section 3.2.3 later on.

To wrap up this subsection on OBMD, let us briefly mention also some related recent methods. A simplified OBMD-like scheme, named particle exchange AdResS (PE-AdResS), has been introduced in Ref. [209]. In that setup, to mimic an open boundary, AdResS is combined with a Metropolis particle exchange criterion to perform an identity swap at eight corners of the simulation domain. The corners act as coarse-grained (for easier particle identity exchange, similarly as in OBMD) reservoirs. Individual solvent components are allowed to fluctuate through the swap of identity of particles based on standard Metropolis criterion, which happens whenever depletion is observed. However, this is only a semi-GC method because the total number of particles in the simulation domain remains constant during the course of simulation,

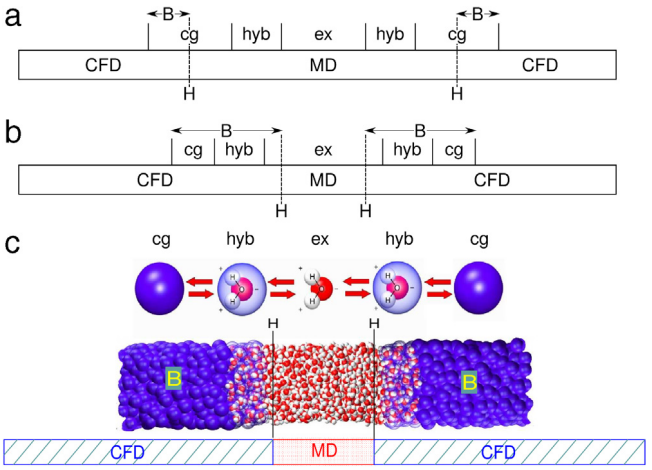

Fig. 10. (c) Domain decomposition of triple-scale liquid water system: coupling of atomistic and continuum hydrodynamics. The MD domain contains atomistic water surrounded by continuum fluid dynamics (CFD) domains in the horizontal direction. Periodic boundary conditions are used in the orthogonal directions. The momentum flux is exchanged across the interface H. (b) Particle buffers B contain a multiscale description, which includes all-atom (*ex*), coarse-grained (*cg*) and hybrid (*hyb*) models. (a) Particle buffers B contain only *cg* description, while the change of molecular resolution takes place inside the MD domain.
Source: Reprinted from Ref. [225].

i.e., particle do not enter or leave the simulation domain, only their identities are swapped. The method was applied to study an experimentally observed coil-globule-coil transition of poly(*N*-isopropylacrylamide) (PNIPAm) in aqueous methanol mixture.

Yet another promising idea has been introduced in Ref. [149], where atomistic liquid models have been coupled to the ideal gas by using both AdResS and H-AdResS. The underlying idea is that employing the ideal gas as the coarse-grained model in the low resolution domain makes intermolecular interactions redundant bringing high computational-efficiency gains. However, the number of molecules in this approach fluctuates only in the high resolution domain by exchanging molecules with the low resolution domain where one imposes the periodic boundary conditions at the boundaries of simulation box in all three directions. On the other hand, the total number of molecules in the system remains constant [210,211]. This is essentially the same as focusing on some subdomain in a closed full-blown atomistic simulation, corresponding to the adaptive resolution simulation where the same atomistic model from the fine-grained domain is used as a coarse-grained model in the low resolution domain (see discussion below Eq. (52)). The variance of fluctuating quantities, e.g., number of molecules in the subdomain of interest, will not be fully in agreement with the GC ensemble (as, on the contrary, they are in OBMD, see Fig. 9) but there will be some finite size effects due to periodic boundary conditions at the boundaries of the simulation box [211]. Of course, in the approach presented in Ref. [149], using the cheapest coarse-grained model in the low resolution domain, i.e., the ideal gas, one can afford to make the low resolution domain very large, mitigating the finite size effects.

3.2.3. Coupling to continuum hydrodynamics

OBMD, described in previous Section 3.2.2, enables us to couple particle-based open MD system with continuum hydrodynamics to describe the external fluid flow. A hybrid method that aims to couple atomistic and continuum descriptions of liquids needs to satisfy the following requirement: physical quantities, e.g., density, momentum, energy, and the corresponding fluxes must be continuous across the interface between the two resolutions where atomistic and continuum domains provide each other with boundary conditions. To impose boundary conditions from the atomistic to continuum domain is non-problematic as it involves temporal and spatial averaging. Imposing the continuum boundary conditions on the particle domain presents the major challenge in development of the hybrid methods. For this purpose, there have been two kinds of schemes presented in the literature: methods that use the state-variable coupling [212–217] and methods based on flux-exchange [206,218,219]. In the following, we will briefly describe, how this is done using OBMD, where Eulerian hydrodynamics and MD are interfaced, and refer the interested reader to reviews on this topic [208,201,220,81].

OBMD allows us to couple an open molecular domain with continuum fluid dynamics domains via the flux-exchange (hybridMD) coupling [200,206,218,221–224], resulting in a triple-scale setup [219,225] depicted in Fig. 10. In this triple-scale approach, the particle-based OBMD domain is simulated by MD simulation. The dynamics of molecules is thus governed by Newton's equations of motion, as explained in Section 3.2.2. On the other hand, continuum description enables the study of macroscopic fluid flows. The fluid is described by Navier–Stokes equation:

$$\rho\left(\frac{\partial \mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}\right)=-\nabla p+\nabla\cdot\boldsymbol{\Pi}+\mathbf{f}, \tag{80}$$

which is derived from the law of conservation of momentum, i.e., macroscopic Newton-like equation of motion. Here, $\mathbf{u}$, p, ρ, and $\mathbf{f}$ represent fluid's velocity field, pressure, density, and external force acting on it, respectively. Stress tensor is given by:

$$\Pi = -\eta[\nabla\mathbf{u}]^S - \xi\nabla\cdot\mathbf{u}\,\mathbf{I}, \tag{81}$$

with dynamic viscosity η and bulk viscosity ξ. The traceless symmetric tensor is defined as $A^S_{\alpha\beta} = (A_{\alpha\beta} + A_{\beta\alpha}) - (2/3)A_{\gamma\gamma}$ [219].

One can recast the Navier–Stokes equation, Eq. (80), in a form of conservation equation as [218]

$$\partial\phi/\partial t = -\nabla\cdot\mathbf{J}^\phi. \tag{82}$$

Here, $\phi(\mathbf{r}, t)$ denotes a conserved fluid variable and $\mathbf{J}^\phi(\mathbf{r}, t)$ the associated local flux. The relevant conservation laws are mass ($\phi = \rho, \mathbf{J}^\phi = \rho\mathbf{u}$), momentum ($\phi = \rho\mathbf{u}, \mathbf{J}^\phi = \mathbf{J}_p = p\,\mathbf{I} + \rho\mathbf{u}\mathbf{u} + \Pi$) (see Eq. (75)), and energy ($\phi = \rho\epsilon, \mathbf{J}^\phi = \mathbf{J}_e = \rho\epsilon\mathbf{u} + \mathbf{J}_p\cdot\mathbf{u} + q$) (see Eq. (76)) ones. In the last conservation law, ϵ denotes energy and q conduction heat flux. Some variables are connected by constitutive relation: equation of state ($p = p(\rho)$), caloric equation of state ($\epsilon = \epsilon(\rho, T)$), stress tensor (Eq. (81)), and conduction heat flux ($q = -k_c\nabla T$).

Eq. (82) can be numerically solved in different ways [226] but we have resorted to finite volume method, where the continuum region is divided into small cells of volume V_C [218,219]. The conservation laws are then integrated over each of the cells:

$$\int_{V_C} \partial\phi/\partial t\, dV = -\int_{V_C} \nabla\cdot\mathbf{J}^\phi\, dV = -\oint_S \mathbf{J}^\phi\cdot dS \tag{83}$$

$$\frac{d\Phi_C}{dt} = -\sum_{f=faces} A_f\mathbf{J}^\phi_f\cdot\mathbf{n}_f. \tag{84}$$

Here, $\Phi_C = \int_{V_C}\phi(\mathbf{r}, t)d\mathbf{r}^3$ and A_f denotes the area of each face that surrounds the volume V_C. $\mathbf{J}^\phi_f$ is calculated as the average of the flux $\mathbf{J}^\phi$ in the two adjacent cells to the face f. The discretized Navier–Stokes equations given by Eq. (84) are then integrated in time using an explicit Euler scheme [224].

The MD and continuum domains share an interface H, as shown in Fig. 10. The otherwise independent MD and continuum domains exchange information after every fixed time interval Δt_c [219]. Flux balance implies the conservation of mass and momentum across H, i.e., both domains should receive equally large but oppositely signed mass and momentum transfer across H over each Δt_c (here, as in Section 3.2.2, for simplicity, we focus only on mass and momentum transfer and leave the explicit energy transfer out of consideration). The momentum flux across the H interface is then used to update the flow variables at the continuum boundary cells, according to Eq. (84). In turn, the same (but oppositely signed) flux needs to be imposed into the particle system across H. This is done using the OBMD protocol described previously (more details can be found in Refs. [81,219,225]). The presented methodology focuses on coupling with nonfluctuating hydrodynamics but it can be also used for the fluctuating case [218,224].

Before turning our attention to bridging to the quantum scale in the next section, let us briefly mention few other recent hybrid approaches interfacing MD and continuum hydrodynamics, which are either based on or adapt a similar coupling strategy as AdResS. In the approach of Petsev at al. [151], AdResS is used to couple MD to smoothed dissipative particle dynamics (SDPD) [227]. SDPD is a particle-based, Lagrangian, continuum solver used to numerically, in an MD-like fashion, solve Navier–Stokes equations. SDPD is a fluctuating extension of smoothed particle hydrodynamics (SPH) [228,229], incorporating thermal fluctuations. The hybrid MD/SPDP approach was applied to Lennard-Jones fluid [151]. On the other hand, Alekseeva et al. [150] applied AdResS to link MD with multiparticle collision dynamics (MPC) [230]. MPC is a mesoscale simulation method for fluid flows, where the fluid is modeled by particles with continuous positions and velocities and stochastic interparticle interactions. The fluid is discretized into cells with no restriction on the number of particles in each cell [230]. MPC, which locally conserves mass, momentum, and energy, models hydrodynamics on large length and times scales. The robustness of the hybrid MD/MPC model was again demonstrated on a Lennard-Jones fluid [150]. Another interesting hybrid approach has been presented in Ref. [231], where the hybrid MD/continuum hydrodynamics system of liquid water was considered as two completely miscible liquids using two-phase modeling, i.e., one phase corresponding to the MD and the other to fluctuating continuum hydrodynamics. Coupling between the two models is achieved by allowing exchange of mass and momentum between the two phases and introducing a parameter that quantifies the distribution of mass and momentum between the phases. This parameter, which varies between 0 and 1, is introduced in a similar fashion as the weighting function in AdResS, Eq. (39), where the value 0 corresponds to MD description whereas 1 describes the continuum phase.

4. Quantum molecular systems with open boundaries: some examples of simulation techniques

The description of quantum systems, as expected, cannot be done via a mere extension of classical theories. Actually, as shown with the model of Emch and Sewell in Section 2.2, it may be easier to consider the classical description as a simpler

subcase of the quantum case. The operator and wavefunction formalism and the peculiarity of the statistics of quantum particles (compared to the classical case) make the treatment of quantum open boundary systems rather challenging. Until today only specific situations and under specific approximations, are treated. In this section, we will discuss representative examples of applications of the concept of exchange of heat and matter between a system and a reservoir for systems described at different levels of quantum accuracy. We will start from the semi-classical (or semi-quantum) approach of path integral/ring polymer representation of atoms in molecular simulation, where the quantum effect of spatial delocalization of atoms is considered. We will pay particular attention to the embedding of the AdResS method within the path integral framework. This allows for the calculation of both static and dynamic properties. Next, we will discuss the extension to an open boundary setup of the so called Quantum Mechanics/Molecular Mechanics (QM/MM) approaches. Here, electrons come into the game. Still in fields related to molecular simulation, we will then show two examples, already cited in the introduction, of techniques which, within the framework of Density Functional Theory (DFT), interface molecular systems with a reservoir of electrons and nuclei. These are the computation of redox potentials which involves electrodes modeled as electrons reservoir [25] and the rational compound design where the sampling of the chemical space is achieved by the GC approach to electrons and nuclei [26,27]. High accuracy in the determination of the electronic properties are assured by wavefunction-based methods. In our context of open boundary systems, we will report about a Hartree–Fock approach for variable number of electrons and this concludes our review of quantum approaches.

4.1. Open boundary systems and path integral approach in molecular simulation

The path integral (PI) formalism of Feynman is a powerful method to describe the quantum character of spatial delocalization of atoms in space at a basic level [232]. In particular, the basic quantum features of systems at low temperature are captured by the PI description [233]; even for systems composed of light atoms at room temperature, e.g., liquid water, due to the presence of hydrogen atoms, the PI treatment is often mandatory [234]. For this reason, quantum effects, due to the spatial delocalization of atoms, are treated in molecular simulation via the path integral technique [47]. The introduction of the PI formalism in molecular simulation/dynamics of molecular systems is a well known approach. Thus, here we describe only the relevant aspects of its formal derivation. More details can be found in Appendix C, while a clear and complete treatment of the subject can be found in, e.g., Refs. [47,235]. The relevant part of the idea for molecular simulation is the transformation of a classical Hamiltonian of N distinguishable particles (atoms) in a quantized Hamiltonian, which is formally equivalent to an effective Hamiltonian of classical polymer rings each of which represents one atom. If one starts from a classical Hamiltonian of atoms with phase space coordinates ($\mathbf{p}$, $\mathbf{r}$), mass m_j (mass of the jth atoms) and interaction potential $V(\mathbf{r}_1, \ldots, \mathbf{r}_N)$:

$$H = \sum_{j=1}^{N} \frac{\mathbf{p}_j^2}{2m_j} + U(\mathbf{r}_1, \ldots, \mathbf{r}_N), \tag{85}$$

then the PI formalism allows to transform Eq. (85) into:

$$H_P = \sum_{i=1}^{P} \left\{ \sum_{j=1}^{N} \frac{1}{2} m_j \omega_P^2 (\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + \frac{1}{P} U(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i) \right\}, \tag{86}$$

where the classical kinetic term for an atom is transformed into an effective polymer ring. Its P beads are linked via a first-neighbor harmonic interaction and the classical interatomic potential is now distributed over the beads in such a way that each bead of a polymer ring interacts with the corresponding bead of another polymer ring (i.e. of another atom). In simple terms, the quantized Hamiltonian corresponds to an isomorphism between a quantum system and a classical system of ring polymers. Here, $\omega_P = \frac{\sqrt{P}}{\beta\hbar}$ ($\beta = 1/k_BT$) and $U(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i)$ is the potential that acts between same bead index i of two different particles (see Fig. 11 for a pictorial illustration). The effect of spatial delocalization of an atom is mimicked by the spatial fluctuation of the polymer ring and by the distribution of the interactions over different sites. The implication is that by sampling configurations in space via molecular simulation of a system composed of "effective" classical polymers one would obtain a basic quantum statistical description of the "real" system. Eq. (86) suggests that a natural approach to sample such configurations for the calculation of static properties can be done via a standard MC technique. However, via a further formal manipulation, it is also possible to employ MD for both static and dynamic properties. The technical aspects of this procedure are reported in detail in Appendix D. Next, we will focus on the description of both cases: (a) the PI approach in the GC MC technique and the PI approach within the MD technique for open boundary systems (in the AdResS fashion).

4.1.1. Path integral Monte Carlo in the grand canonical ensemble

The quantum–classical isomorphism of Eq. (86) reduces the GC MC sampling of a quantum system in real space to a standard GC MC of a classical system of ring polymers. Hence, the application of this technique is (essentially) very similar to that already seen for classical systems in contact with a particle reservoir, in Section 3. In this perspective, the paper of Wang, Johnson and Broughton [236] is considered the reference work for GC MC in real-space path integral simulations. In the following, we will report an account of their method.

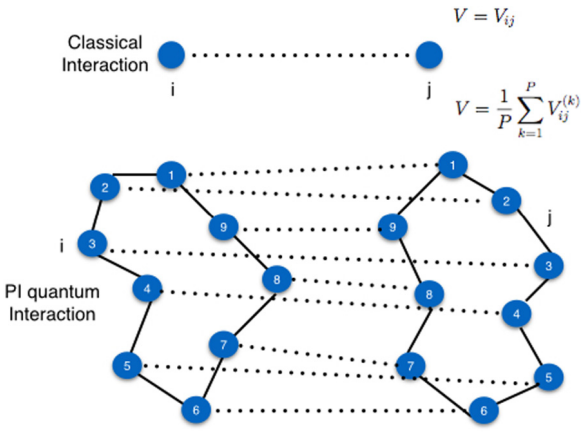

Fig. 11. Pictorial illustration of the quantization of classical atoms due to the path integral formalism via the mapping to a system of classical ring polymers.

The GC partition function is written as:

$$\Xi(\mu, V, T) = \sum_N e^{\beta\mu N} Q(N, V, T), \tag{87}$$

where μ is the chemical potential and $Q(N, V, T)$ the canonical partition function, which in the PI formalism (see Appendix C) is written as: $Q = \frac{1}{N!}\left(\frac{2\pi mPkT}{h^2}\right)^{3NP/2} \int \exp\left[-\beta\sum_{i=1}^{P}\left\{\sum_{j=1}^{N}\frac{1}{2}m\omega_P^2(\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + \frac{1}{P}U(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i)\right\}\right] d\mathbf{r}_1 \ldots d\mathbf{r}_N$ (for simplicity $m_j = m$; $\forall j$). It follows that:

$$\Xi(\mu, V, T) = \sum_N e^{\beta\mu N}\frac{1}{N!}\left(\frac{2\pi mPkT}{h^2}\right)^{3NP/2} \times \int \exp\left[-\beta\sum_{i=1}^{P}\left\{\sum_{j=1}^{N}\frac{1}{2}m\omega_P^2(\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + \frac{1}{P}U(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i)\right\}\right] d\mathbf{r}_1 \ldots d\mathbf{r}_N. \tag{88}$$

The probability density is defined as:

$$\mathcal{P}(N) = \frac{1}{\Xi}\frac{V^N e^{\beta\mu N}}{N!}\left(\frac{2\pi mPkT}{h^2}\right)^{3NP/2} \times \exp\left[-\beta\sum_{i=1}^{P}\left\{\sum_{j=1}^{N}\frac{1}{2}m\omega_P^2(\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + \frac{1}{P}U(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i)\right\}\right]_N. \tag{89}$$

The displacements of the molecules are driven by the trial move accepted with probability: $\mathcal{P}^{disp} = min(1, \exp(-\beta\Delta H_P))$, where ΔH_P = is the change in total energy. The crucial move, compared to the MC moves for systems with fixed N, is the creation and deletion of molecules. To this aim, let us consider a system with N molecules in a certain state S_N, next a new configuration $S_{N+1|\Gamma}$ is generated by inserting a molecule in a certain configuration Γ. It follows that the probability of observing the new configuration is: $G(N \to N+1|\Gamma) = N! \times \mathcal{P}(N) \times \mathcal{P}(N \to N+1|\Gamma) \times \mathcal{P}^{acc}(N \to N+1|\Gamma)$. Here, the factor $N!$ corresponds to the counting of all possible permutations of the molecules, $\mathcal{P}(N)$ is the probability that the system is in S_N, $\mathcal{P}(N \to N+1|\Gamma)$ is the probability of choosing a specific conformation Γ out of all possible conformations of the ring polymers and finally $\mathcal{P}^{acc}(N \to N+1|\Gamma)$ is the probability of acceptance. The reverse move is driven by the probability: $G(N+1|\Gamma \to N) = (N+1)! \times \mathcal{P}(N+1|\Gamma) \times \mathcal{P}(N+1|\Gamma \to N) \times \mathcal{P}^{acc}(N+1|\Gamma \to N)$. The acceptance criterion for a (reversible) particle creation is:

$$\mathcal{P}^{create} = min\left(1, \frac{\mathcal{P}^{acc}(N \to N+1|\Gamma)}{\mathcal{P}^{acc}(N+1|\Gamma \to N)}\right). \tag{90}$$

Detailed balance is accomplished equalizing the probability of the forward move and that of reverse move: $G(N \to N+1|\Gamma) = G(N+1|\Gamma \to N)$, from which it follows:

$$\frac{\mathcal{P}^{acc}(N \to N+1|\Gamma)}{\mathcal{P}^{acc}(N+1|\Gamma \to N)} = \frac{(N+1)! \times \mathcal{P}(N+1|\Gamma) \times \mathcal{P}(N+1|\Gamma \to N)}{N! \times \mathcal{P}(N) \times \mathcal{P}(N \to N+1|\Gamma)}. \tag{91}$$

Given the definition of $\mathcal{P}(N)$ in Eq. (89), the ratio: $\frac{\mathcal{P}(N+1|\Gamma)}{\mathcal{P}(N)}$ essentially isolates the inserted particle in configuration Γ, as a consequence we have:

$$\frac{\mathcal{P}^{acc}(N\to N+1|\Gamma)}{\mathcal{P}^{acc}(N+1|\Gamma\to N)} = Ve^{\beta\mu}\exp\left[-\beta\left(\sum_{i=1}^{P}\left\{\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2+\sum_{j=1}^{N}\frac{1}{P}U(\mathbf{r}_{j\gamma}^{i})\right\}\right)\right] \times\left(\frac{2\pi mPkT}{h^2}\right)^{3NP/2}\frac{\mathcal{P}(N+1|\Gamma\to N)}{\mathcal{P}(N\to N+1|\Gamma)}, \tag{92}$$

where $\mathbf{r}_\gamma^{(i)}$ indicates the ith bead of the inserted molecule γ (with conformation Γ), $\mathbf{r}_{j\gamma}^{i}$ is the bead–bead distance between the bead i of molecule γ with the bead i of the molecule j, and $\sum_{j=1}^{N}\frac{1}{P}U(\mathbf{r}_{j\gamma}^{i})$ is the total interaction energy of molecule γ with the other N molecules in state S_N. In this approach, the sampling of conformations of the inserted molecule is done as ideally picking it up from a gas of ring polymer fluid, that is the probability of choosing conformation Γ, $\mathcal{P}(N\to N+1|\Gamma)=\mathcal{P}(\Gamma)=\frac{\exp\left[-\beta\left(\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2\right)\right]}{\int d\Gamma\exp\left[-\beta\left(\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2\right)\right]}$. Moreover, one has a probability $\frac{1}{N+1}$ of choosing the molecule with conformation Γ from a state with $N+1$ molecules in which at least one must have conformation Γ, that is $\mathcal{P}(N+1|\Gamma\to N)=\frac{1}{N+1}$. From the considerations above it follows that:

$$\frac{\mathcal{P}^{acc}(N\to N+1|\Gamma)}{\mathcal{P}^{acc}(N+1|\Gamma\to N)} = \frac{V}{N+1}e^{\beta\mu}\exp\left[-\beta\left(\sum_{i=1}^{P}\left\{\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2+\sum_{j=1}^{N}\frac{1}{P}U(\mathbf{r}_{j\gamma}^{i})\right\}\right)\right] \times\left(\frac{2\pi mPkT}{h^2}\right)^{3NP/2}\frac{\int d\Gamma\exp\left[-\beta\left(\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2\right)\right]}{\exp\left[-\beta\left(\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2\right)\right]}. \tag{93}$$

Taking into account that the partition function of a single quantum particle is nothing else but the partition function of a free particle and assuming that we deal with a structureless particle, it follows that the partition function of an isolated path (ring polymer with P beads) is: $\int d\Gamma\exp\left[-\beta\left(\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(\mathbf{r}_\gamma^{(i)}-\mathbf{r}_\gamma^{(i+1)})^2\right)\right]=\frac{(2\pi mkT)^{3/2}}{(2\pi mPkT)^{3P/2}}$. Thus, we obtain:

$$\frac{\mathcal{P}^{acc}(N\to N+1|\Gamma)}{\mathcal{P}^{acc}(N+1|\Gamma\to N)} = \frac{V\exp\left(-\beta\left[\sum_{j=1}^{N}\frac{1}{P}U(\mathbf{r}_{j\gamma}^{i})\right]\right)}{N+1}\left(\frac{2\pi mkT}{h^2}\right)^{3/2}e^{\beta\mu}. \tag{94}$$

Considering further that $\Lambda=\left(\frac{2\pi mkT}{h^2}\right)^{-1/2}$ is the thermal wavelength, we finally obtain:

$$\mathcal{P}^{create}=min\left(1,\frac{V\exp[\beta(\mu-\sum_{j=1}^{N}\frac{1}{P}U(\mathbf{r}_{j\gamma}^{i}))]}{(N+1)\Lambda^3}\right). \tag{95}$$

Regarding $\mathcal{P}^{delete}$, that is the probability of removing a particle, one can think in the following terms: from a configuration with N particles a new configuration with $N-1$ particles is generated by choosing a molecule in a random way and removing it. The (microscopically) reversible step would have been made by inserting a molecule with the configuration of the molecule deleted in a system with $N-1$ particles. Given the condition of detailed balance applied to the insertion of a molecule in a system with $N-1$ particles, we can then write:

$$\mathcal{P}^{delete}=min\left(1,\frac{N\Lambda^3}{V}\exp\left[\beta\left(\sum_{j=1}^{N-1}\frac{1}{P}U(\mathbf{r}_{j\gamma}^{i})-\mu\right)\right]\right), \tag{96}$$

where Γ is the conformation of the deleted molecule and the interaction energy corresponds to that of the deleted molecule with the remaining $N-1$ molecules. This approach has been applied, as test case, to liquid parahydrogen and liquid neon and the comparison of the simulation results with experimental data is highly satisfactory. The method, reported here, represents a seminal work for the GC MC treatment of molecules in the path integral formalism, further work was later on done along this direction where technical improvements to the basic algorithm were introduced (see e.g. [237] and references therein). Dynamical properties, instead, cannot be calculated with such methodology in a straightforward way. In this perspective, in the next section, we will introduce the path integral methodology in the GC fashion performed through MD. Such an approach can be realized via the (GC-)AdResS technique, as done for classical systems, adapted to the various PI techniques of MD.

4.1.2. Path integral molecular dynamics in the adaptive resolution Grand Canonical fashion

The original idea of merging PIMD and AdResS was based on a simple extension of the AdResS principle. The dynamics of polymer rings, from a technical point of view, is nothing else but the dynamics of classical DOFs. Hence, the standard AdResS

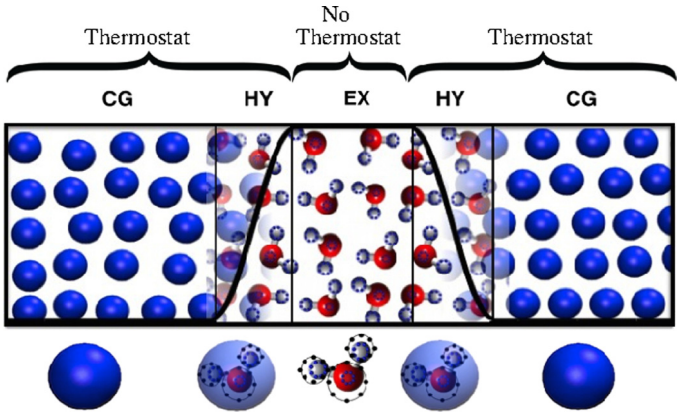

Fig. 12. Pictorial representation of the GC-AdResS scheme. CG indicates the coarse-grained region, HY the hybrid region where path-integral and coarse-grained forces are interpolated via a space-dependent, slowly varying, function $w(x)$ and the explicit resolution region, EX (or PI), is the path-integral region (that is the region of interest). As an example, we report the case, where the "local" thermostat technique is employed so that the dynamical properties in the EX region are not altered by the action of the thermostat.
Source: Figure adapted from Ref. [241].

could be applied (technically) in the same way [238–240]:

$$F_{\alpha\beta} = w_\alpha w_\beta F^{PI}_{\alpha\beta} + [1 - w_\alpha w_\beta] F^{CG}_{\alpha\beta}, \tag{97}$$

where the only modification/extension is the presence of $F^{PI}_{\alpha\beta}$, that is the force between beads of the rings representing the atoms of molecule α and molecule β, instead of $F^{AT}_{\alpha\beta}$ (see also Fig. 12 for a pictorial representation). However, the force-based approach does not allow, from the conceptual point of view, a rigorous Hamiltonian coupling between the ring polymers and the coarse-grained molecules. Since a rigorous Hamiltonian formalism is mandatory for the PIMD approach, the initial formulation of the problem, from the conceptual point, as done in Refs. [238–240], was based on a mere empirical intuition. Nevertheless, the results of the simulation were numerically satisfactory. Later on, the formalization of AdResS in GC-AdResS, reported in the previous section, provided, as for the classical case, a solid theoretical basis to the setup of PI-AdResS. The basic principle is based on the same arguments employed for the classical GC-AdResS, that is the coupling force term between the ring polymers and the hybrid/coarse-grained molecules of the hybrid/coarse-grained region is conceptually not relevant as long as the reservoir (hybrid plus coarse-grained region) conserves the macroscopic quantity (temperature and density). However, as for the classical case, one can actually explicitly calculate the energetic contribution corresponding to the coupling force between the EX region and the reservoir (EX-Res) and show that this contribution is negligible compared to the interaction energy of the molecules of the EX region only. The total energy of coupling at time t, by simple extension to the polymer ring formalism of Eq. (72) in Section 3.2.1 is:

$$W_{EX-Res}(t) = \sum_{i \in EX} W^i_{EX-Res}(t)\,. \tag{98}$$

$W_{EX-Res}(t)$ must be compared for any t with the amount of energy, $W_{EX-EX}(t)$, that is the energy corresponding to the interaction between molecules of the EX region only: $W_{EX=EX}(t) = \sum_{i<j} U^{ij}_{EX}$; $i, j \in EX$. If

$$\frac{|W_{EX-EX}(t)| - |W_{EX-Res}(t)|}{|W_{EX-EX}(t)|} \approx 1; \forall t \tag{99}$$

then the total energy of the EX region up to a negligible correction corresponds to the energy associated with the Hamiltonian of the EX region only. This allows to write a well defined and accurate quantized Hamiltonian for PI-AdResS:

$$\mathcal{H}^M_P = \sum_{i=1}^{P} \left[\sum_{j=1}^{M} m_j \omega_P^2 (\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + \frac{1}{P} U(\mathbf{r}^i_1, \ldots, \mathbf{r}^i_M) \right]; \quad \{\mathbf{r}_1, \ldots, \mathbf{r}_M\} \in EX \tag{100}$$

where M is the instantaneous (variable) number of molecules in the EX region.

Fig. 13 shows that condition (99) holds in all simulations done so far, even in those where the technical conditions are not optimal (i.e. the size of the EX region much smaller that usually considered in AdResS simulations) [241,242]. The GC-AdResS method was implemented in different PIMD schemes such as two versions of Ring Polymer Molecular Dynamics (RPMD) named here $\mathbf{H}_2$ and $\mathbf{H}_3$ (see Appendix D.0.1 for details) and Centroid Molecular Dynamics (CMD) (see Appendix D.0.2 for details). Calculations of static and dynamic properties for various systems were performed [241,242]. Here, as an illustrative example of the robustness of the approach, we report representative results for liquid water at room condition obtained with

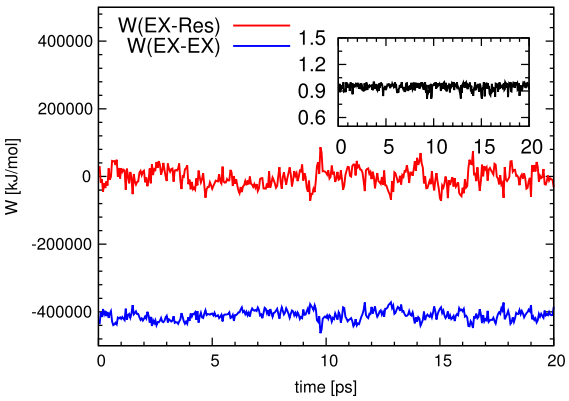

Fig. 13. Main figure: $W_{EX-EX}(t)$ compared to $W_{EX-Res}(t)$. Inset: The relative amount of the interaction between the EX region and the rest of the system along the trajectory: $\frac{|W_{EX-EX}(t)|-|W_{EX-Res}(t)|}{|W_{EX-EX}(t)|}$; the contribution is, at most, of 10%. The example reported here refers to the calculation done with RMPD, however also for the other approaches one obtains the same results.
Source: Figure adapted from Ref. [241].

the RPMD techniques named $\mathbf{H}_2$, $\mathbf{H}_3$ and with the CMD technique; see Fig. 14 (static properties), Figs. 15 and 16 (dynamic properties). As a matter of fact, PI-GC-AdResS represents the first approach of PIMD in a GC (open boundary) fashion. It must be reported that recently, the PI formalism has been also introduced in the H-AdResS version. In that example, the technical novelty of a space-dependent particle mass [243] was introduced. This technical development allows to pass from a standard polymer ring in the EX region to a collapsed ring (one interaction site) in the coarse-grained region. This scheme, as long as the hypothesis of negligible derivative of the mass as a function of space holds, has been shown to be robust when applied to a simplified test-case of liquid parahydrogen [194]. However, this approach has not been tested yet on more realistic chemiophysical systems and the physical consequences of an artificial, non-physical definition of atomic space dependent mass have not been explored in full, yet.

4.2. QM/MM with open boundaries

An interesting idea, that since its appearance in 1976 [244] has led to important breakthroughs (above all, but not only) in biochemistry, is the molecular simulation technique named QM/MM [245–247]. The idea is that of interfacing a region of the system treated at quantum mechanical level (QM) with a (usually) larger region treated at the level of classical molecular mechanics (MM). The QM region is characterized by properties that require a resolution at electronic level. The basic idea is that the overall role of the environment, although often not dominant, is of high importance and thus it is mandatory to take it into account. A fully quantum treatment of the system would be prohibitive. Under the hypothesis that the electronic properties of the large environment around a certain region are not relevant, the treatment of its molecules can be done at classical level, i.e., the environment plays the role of an effective particle-based thermodynamic bath. A typical example is that of a chemical reaction in complex physicochemical systems such as solutions [248] where the reactants are localized in space and the reaction mechanism involves only their electrons. However, the reactivity is strongly affected by the overall (thermodynamic) action of the solvent on the reactants. The key aspect of the QM/MM method is the coupling between the two regions. The original scheme considers the QM and the MM region fixed regarding the number of particle in each region (i.e. there is no particle exchange). The progress in time in engineering the interface/coupling setup has brought QM/MM methods of the last generation into the category of method which can treat open boundary systems with variable number of molecules [170,249–254]. It must be clarified that the explicit intention of the developers is of mere technical character, that is of minimizing the computational errors in the QM part due to the coupling to the MM part. In this sense, the adaptive QM/MM, which can be grouped into the category of open boundary approach, shall not be considered a systematic attempt to build a first-principle procedure for an open boundary quantum system with a reservoir of molecules composed of nuclei and electrons. However, at least at technical level, these methods represent a pragmatic progress towards the treatment of open boundary systems at quantum level for relevant physicochemical problems. Specific technical details about the general QM/MM approach are given in Appendix E, while here we report only the basic information required to understand the approach of open system (i.e. QM system with varying number of molecules) which emerges in adaptive QM/MM. The primary interest of the developers of adaptive QM/MM is that of having active sites of a systems, e.g., reactive chemical units of a molecule as binding sites for solvent molecules or other molecules. Around each active site, the interaction is treated at quantum mechanical level by considering interacting molecules as quantum objects if they are within a certain distance from the active site or as classical objects if they are outside a certain region. Moreover, molecules can diffuse and change their identity from classical to quantum as the system evolves in time. In this sense, strictly speaking, the method has more the character of multiscale approach rather than of a typical (physical) open boundary system. However, for our current

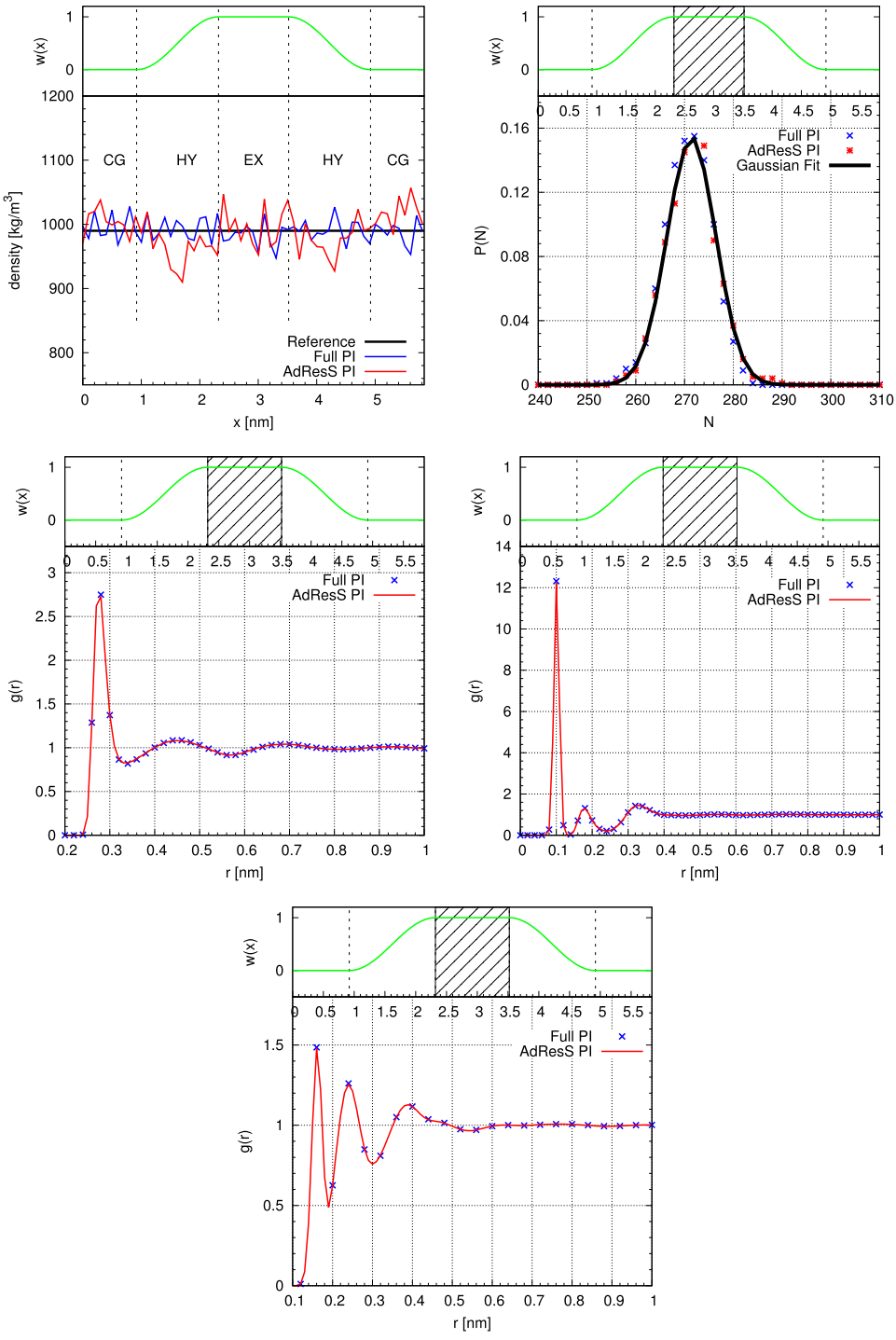

Fig. 14. From left to right (top): Particle number density and particle number probability distribution. From left to right (bottom): (bead–bead) oxygen–oxygen, oxygen–hydrogen and hydrogen–hydrogen partial radial distribution functions. Such functions are compared with the results obtained for an equivalent subsystem ($EX = 1.2$ nm) in a full path integral (Full PI) simulation. These results were obtained by the RMPD technique $\mathbf{H}_3$. However, we obtain results with the same satisfactory agreement also using the other PI techniques implemented in AdResS.
Source: Figure adapted from Ref. [241].

interest, we can imagine to have a system with one "active" site around which a QM region is defined and is interfaced with a reservoir of energy and particles, i.e. the MM region. The major problem in such schemes is that the fluctuations of number of molecules in the QM region imply a drastic change of the total energy of the system. In order to minimize this

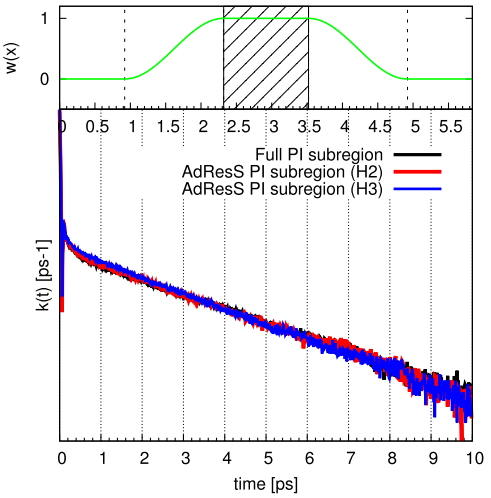

Fig. 15. The rate function $k(t)$ for q-SPC/FW water model calculated in quantum subregion of GC-AdResS and an equivalent subregion in RPMD ($\mathbf{H}_3$ and $\mathbf{H}_2$) simulations. $k(t)$ is the average rate of change of hydrogen-bond population in trajectories where the bond is broken at after a time t (from an initial reference time). $k(t)$ is derived from the correlation function: $c(t) = \langle h(0)h(t)\rangle / \langle h\rangle$, where $h(t)$ is the hydrogen bond population operator, with value 1, for bonded pairs, and zero otherwise. Then $k(t) = -dc/dt$.
Source: Figure adapted from Ref. [241].

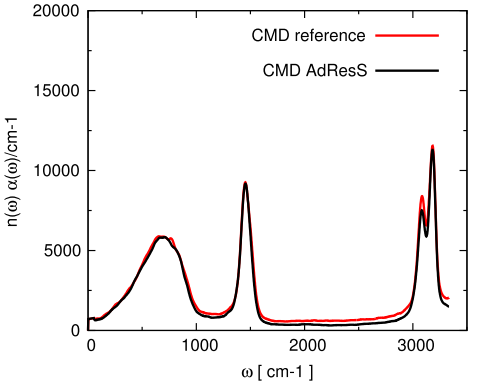

Fig. 16. Infrared spectrum for liquid water at 298 K calculated in explicit region of AdResS CMD and an equivalent subregion in the reference CMD simulations.
Source: Figure adapted from Ref. [242].

problem a buffer or transition region, similar to that introduced by classical and path integral adaptive resolution methods, is introduced (see Fig. 17 for a schematic representation). Differently from the classical or semiclassical/path integral concept of adaptivity, in this case one must deal with electrons. A paradox then emerges, that is a molecule in the hybrid region would have a fractional quantum character or, due to the delocalization of electrons, part of the electron cloud is in one resolution and part in another. For a quantum mechanical treatment, one needs that either molecules are treated at full quantum level, or as in the standard QM/MM scheme with rigid boundaries, as classical molecules whose role within the QM Hamiltonian is that of providing external interactions. This problem is solved with a principle common to most of the current adaptive QM/MM methods, that is at each timestep of the simulation, the buffer region is partitioned in different subsets. Next, the (standard) QM/MM potential is defined for all the possible partitionings and, for each partitioning, the molecules of the corresponding subset of the buffer are included in the QM region. Thus, for each of these "extended" QM regions a standard QM/MM calculation is done. The total potential is then defined as a weighted average of these individual potentials: $U(\mathbf{r}) = \sum_i^M f_i(\mathbf{r})U_i(\mathbf{r})$, where each $U_i(\mathbf{r})$ corresponds to one of the M partitioning of the system in a group of QM molecules and a group of MM molecules and $f_i(\mathbf{r})$ is the weighting function. The weighting function $f_i(\mathbf{r})$ is function of the coordinates of the molecules and can be expressed in terms of single molecule switching functions. The switching function is constructed to follow the general principle that the quantum energy of molecules far way from the active site counts less that the energy of those which are closer. Next, the dynamical evolution, based on such potential energy, is performed and it creates a new configuration on which the partitioning step is applied once again and the procedure for the calculation of the potential energy follows as described before. Such an idea has been successfully implemented by several groups, with

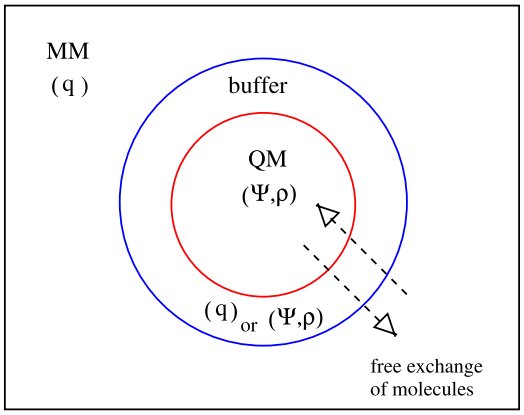

Fig. 17. Schematic representation of the adaptive QM/MM setup. Molecules in the QM region are always treated at quantum mechanical level and thus their electronic properties are calculated explicitly in terms of the wavefunction Ψ or of the electron density $\rho(\mathbf{r})$. In the buffer region, molecules are treated at both levels classically with an effective localized Coulomb charge q at the interaction sites, or quantum. The treatment of a molecules depends on, whether it belongs to the QM or MM subset of a particular partitioning of the buffer zone. Finally, the average is made over all possible partitionings of the buffer zone.

specific choices of the switching function, of the partitioning scheme [249–252] or based on a force interpolation rather than an energy interpolation [253,254]. Despite numerical results are encouraging, from the technical point of view, due to the partitioning procedure, the number of QM calculations required (compared to standard QM/MM studies) may represent an expensive effort. Furthermore, at conceptual level the theoretical framework as open boundary approach is not yet solid. As underlined before, the process of variation of number of molecules in the QM region is achieved via an artificial path, in the sense that the introduction of molecules in the QM region does not follow or satisfy rigorous principles of statistical mechanics. For example, it has not be shown that nuclei and electrons are introduced into the QM region according to a chemical potential (for the nuclei and for the electrons) corresponding to the thermodynamic conditions expected, as instead it should be for a GC-like setup of system with nuclei and electrons. Hence, the thermodynamic conditions, in which the QM region effectively is, may not correspond to the conditions expected. In turn, this implies that one should always check, case by case, that the adaptive QM/MM study reproduces some reference results (from experiment or larger QM/MM). In this sense, the current technical setups of adaptive QM/MM methods cannot be taken *a priori* as predictive tool of GC systems, however, they provide for sure a very solid technical first step towards a rigorous framework to treat the dynamical evolution of electrons and nuclei in the GC fashion. In the next section, we consider techniques, where the GC setup is explicitly defined through the introduction of nuclei and electrons according to the corresponding chemical potential. Compared to QM/MM studies, in these methods only static properties are considered and thus there is no advantage of dynamical evolution of the system, as offered by the QM/MM approach.

4.3. Density functional theory with a particle reservoir

The development of DFT [255,256] has brought a revolution in the study of condensed matter systems at quantum level. The theory is based on the Hohenberg–Kohn (HK) theorem [257] which, in essence, provides a one-to-one correspondence between the ground state properties of a system and its corresponding one-particle electron density. The groundbreaking consequence is that the explicit (and often prohibitive) calculation of the $3N$-dimensional wavefunction of a N-electron system is no more required in order to know its ground state properties. It is actually enough the knowledge of a 3-dimensional quantity, i.e., the one-particle electron density. The formal principle was then made pragmatic by further (technical) simplifications of the idea in terms of Kohn–Sham orbitals [258], that is the most used approach to do electronic structure calculations. In its traditional form DFT was developed for systems with fixed number of electrons and nuclei. However, an electronic GC theory, i.e., system with a reservoir of electrons, was established by the extension of the HK energy functional to the so-called Mermin functional [259]. Furthermore, the attempt to go beyond the Born–Oppenheimer approximations (i.e. the separation of electronic and nuclear Hamiltonian) brought to a formalization of a molecular GC theory within DFT [260], referred in some work as the CNP theory [26]. In this section, we describe the essence of the Mermin's and CNP idea and discuss two important extensions/applications. The basic principles and formulas of DFT are summarized in Appendix F. Here, we report only those aspects required for the discussion of DFT in terms of open system/GC setup. The Mermin functional was developed in order to extend the Hohenberg–Kohn theorem beyond the ground state to nonzero temperatures. In the HK theory the energy functional corresponding to the electronic Hamiltonian, H_{el} with external potential $v(\mathbf{r})$ ($H_{el} = \hat{T} + \hat{V}_{ee} + v(\mathbf{r})$) is : $E[\rho] = \mathcal{F}[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}$, with $\mathcal{F}[\rho]$ the universal HK functional common to all electron systems. The minimum of $E[\rho]$ w.r.t. $\rho(\mathbf{r})$ yields the density of ground state $\rho_0(\mathbf{r})$. For a nonzero temperature T and at a given chemical potential μ, Mermin derived the following energy functional (Grand Potential):

$\Omega[\rho](\mu, T) = \mathcal{F}_{gc}[\rho](\mu, T) - \mu N + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r}$ (see Appendix F). He has shown that its minimum w.r.t. ρ yields the electron density of equilibrium $\rho_0(\mathbf{r})$ at the given thermodynamic conditions, (T, μ), in the presence of the external potential $v(\mathbf{r})$ and $\rho_0(\mathbf{r})$ is the equilibrium density of an electronic GC ensemble. The limit to $0K$ can also be derived (i.e. electronic ground state in the GC ensemble) and the electronic chemical potential, besides the thermodynamic analogy with classical statistical mechanics, is interpreted as an indicator of the escaping tendency of electronic cloud or, more precisely, is approximately equal to the (minus of the) electronegativity [256]. There are few conceptual drawbacks that have been subject of intense discussions such as the justification of the concept of fractional number of electrons in DFT and the differentiability of functionals w.r.t. the number of electrons. In fact, being an electron not localized in space, part of it may be out of the system considered, thus *"fractional electron number may arise as a time average in an open system"* [261]. Despite these questions have been treated intensely they still remain matter of concern. However, satisfactory answers/solutions have been proposed so that the idea of electronic GC (KS)-DFT is solid enough to be applied to the study of various systems [261–263] (see also Appendix F). Tavernelli, Vuilleumier and Sprik, provided an example of important application of the idea of electronic GC approach, that is, the study of the redox process where the electrodes of a concentration cell is replaced by a reservoir of electrons [25]. In such case, the problem of fractional electron number has been avoided by considering two different energy surfaces each corresponding to a system with an integer number of electrons. In such a way, the Grand Potential at a chemical potential μ and temperature T is defined as: $\Omega_{el}(\mathbf{R}) = -k_B T \ln[z^{N_0-1} e^{-\beta E_+(\mathbf{R})} + z^{N_0} e^{-\beta E_0(\mathbf{R})}]$, with $z = \exp(\beta\mu)$ and $\mathbf{R}$ the nuclei configuration in space, N_0 is the number of electrons of a neutral molecule to which corresponds a surface energy $E_0(\mathbf{R})$, while the first cation has number of electrons equal to $N_0 - 1$ and corresponding energy surface $E_+(\mathbf{r})$. The GC scheme is then embedded into a molecular dynamics procedure by defining the force acting on the Ith atom as: $\mathbf{F}_I = -\frac{\partial \Omega_{el}}{\partial \mathbf{R}_I}$. The electronic optimization is performed separately for the N_0 and the $N_0 - 1$ for a given nuclei configuration. So far, we have seen the case where the GC framework of DFT considers only an electron's reservoir. However, a step forward was done by considering the case of molecular Grand Ensemble within the DFT theory. The seminal work of Capitani, Nalewajski and Parr (CNP) [260] was originally developed to take into account non-Born–Oppenheimer (N-BO) Hamiltonians in DFT, however the formalism paved the way to the idea of GC-like treatment of electrons and nuclei. In the N-BO approach, the electronic and nuclear DOFs are no more separated so that the Hamiltonian of N electrons and M nuclei (in atomic units) is:

$$H = -\frac{1}{2}\sum_{i=1,N} \nabla_i^2 - \sum_j -\frac{1}{2M_j}\nabla_J^2 - \sum_{i=1,N}\sum_{j=1,M} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} + \sum_{i<s}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_s|} + \sum_{j<l}^{M} \frac{Z_j Z_l}{|\mathbf{R}_j - \mathbf{R}_l|}, \tag{101}$$

with M_j and Z_j the mass of the charge of the jth nucleus. The corresponding ground state wavefunction, $\Psi(\mathbf{r}_1.....\mathbf{r}_N, \mathbf{R}_1, \ldots, \mathbf{R}_M)$, solution of the Schrödinger equation: $H\Psi = E\Psi$ can then be used to define the one particle electron density:

$$\rho(\mathbf{r}) = \rho_e(\mathbf{r}_1) = N \int \Psi\Psi^+ d\mathbf{r}_2.....d\mathbf{r}_N d\mathbf{R}_1.....d\mathbf{R}_M, \tag{102}$$

and the one particle nuclear density for the nucleus of type a:

$$\rho^a(\mathbf{R}^a) = \rho^a(\mathbf{R}^a)_1 = M_a \int \Psi\Psi^+ d\mathbf{r}_1.....d\mathbf{r}_N d^{(M-1)_a}\mathbf{R}, \tag{103}$$

where $d^{(M-1)_a}\mathbf{R}$ indicates the integration over all the nuclear coordinates of each nucleus species, except $\mathbf{R}_1^a$, and M_a is the number of nuclei of species a. Next, in the spirit of electronic DFT a N-BO ground state density functional is defined as

$$E[\rho_e, \{\rho^a\}] = min\langle \Psi_{\rho_e,\{\rho^a\}} | H | \Psi_{\rho_e,\{\rho^a\}} \rangle. \tag{104}$$

This functional searches over all the wavefunctions with the appropriate symmetry, which integrate to a specific ρ_e and $\{\rho^a\}$ and then delivers the minimum associated energy. Next, in deriving the Euler equation for the ground state density, they define an auxiliary variational functional $\Omega[\rho_e, \{\rho^a\}]$ whose formalism leads to the (effective) treatment of electrons and nuclei in the Grand Ensemble. The auxiliary functional is defined as

$$\Omega[\rho_e, \{\rho^a\}] = E[\rho_e, \{\rho^a\}] - \mu\left[\int \rho_e(\mathbf{r}) d\mathbf{r} - N\right] - \sum_{\alpha=a,s} \lambda_\alpha \left[\int \rho^\alpha(\mathbf{R}^\alpha) d\mathbf{R}^\alpha - M_\alpha\right], \tag{105}$$

where the index α covers all the nucleus species, and μ and λ_α are Lagrange multipliers. The solution of the set of associated Euler equations

$$\frac{\delta\Omega}{\delta\rho_e(\mathbf{r})} = \frac{\delta E}{\delta\rho_e(\mathbf{r})} - \mu = 0, \tag{106}$$

and

$$\frac{\delta\Omega}{\delta\rho^\alpha(\mathbf{R}^\alpha)} = \frac{\delta E}{\delta\rho^\alpha(\mathbf{R}^\alpha)} - \lambda_\alpha = 0, \tag{107}$$

leads to the ground state densities: $\rho_e(\mathbf{r})^{G.S.}$; $\rho^\alpha(\mathbf{R}^\alpha)^{G.S.}$. Importantly, μ and λ_α can be interpreted as chemical potentials of the respective components and such identification is of key importance for the use of such ideas to construct a molecular GC

ensemble (i.e. simulations at constant chemical potential and variable number of electrons and nuclei, see also Appendix F and Ref. [263]). The Euler equations (106), (107) allow to derive the definition of thermodynamic quantities which involve the variations of the number of electrons and nuclei, as it is reported in the example below. As an example, the basic framework outlined above was taken as a basis to develop a molecular GC DFT approach for exploring the chemical space in search of novel molecules with optimized properties. The method has been proposed by von Lilienfeld and Tuckerman [26] and is developed within the general framework of the alchemical free energy perturbation method discussed in Section 3.1.1; they employ a functional similar to (inspired by) that of CNP. However, they also conclude that at zero temperature and when the nuclei are treated at classical level (i.e. strictly localized in space) and so the nuclear density becomes: $Z(\mathbf{r}) = \sum_I N_I \delta(\mathbf{r} - \mathbf{R}_I)$, ($N_I$ is the total number of protons in atom I) instead of a global multiplier λ_α corresponding to the nucleus species α as in the CNP, it is necessary a local multiplier of the form $\lambda(\mathbf{r})$. A modification of the CNP auxiliary functional is then proposed:

$$\Omega[N_e; \rho, Z] = E[N_e; \rho, Z] - \mu_e \left(\int \rho(\mathbf{r}) d\mathbf{r} - N \right) - \int d\mathbf{r} \mu_n \left(Z(\mathbf{r}) - \sum_I N_I \delta(\mathbf{r} - \mathbf{r}_I) \right), \tag{108}$$

where N_e, ρ, μ_e are the total number of electrons, the electron density and the electron chemical potential respectively, and $\mu_n(\mathbf{r})$ is the position dependent Lagrange multiplier associated to the nuclear density. The associated Euler equations, as for CNP, determine the exact ground state densities $\rho(\mathbf{r})$ and $Z(\mathbf{r})$, the stationarity of Ω leads to the conditions:

$$\mu_e = \left[\frac{\delta E[N_e; \rho, Z]}{\delta \rho(\mathbf{r})} \right]_{Z(\mathbf{r})} = \left[\frac{\delta E[N_e; \rho, Z]}{\delta N_e} \right]_{Z(\mathbf{r})}, \tag{109}$$

and

$$\mu_n(\mathbf{r}) = \left[\frac{\delta E[N_e; \rho, Z]}{\delta Z(\mathbf{r})} \right]_{N_e}. \tag{110}$$

The sampling of the chemical space is then achieved by linking different points A and B in chemical space via Kirkwood/adiabatic integration [264]

$$\Delta F = \int_A^B d\lambda \left\langle \frac{\partial E(\lambda)}{\partial \lambda} \right\rangle_\lambda = \int_{N_e^A}^{N_e^B} dN_e \left\langle \frac{\delta E[N_e; \rho, Z]}{\delta \rho(\mathbf{r})} \right\rangle_{N_e} + \int d\mathbf{r} \int_{Z^A(\mathbf{r})}^{Z^B(\mathbf{r})} \left\langle \frac{\delta E[N_e; \rho, Z]}{\delta Z(\mathbf{r})} \right\rangle_{Z(\mathbf{r})}, \tag{111}$$

which corresponds to

$$\Delta F = \int_{N_e^A}^{N_e^B} dN_e \langle \mu_e \rangle_{N_e} + \sum_I \int_{Z^A(\mathbf{r})}^{Z^B(\mathbf{r})} dZ(\mathbf{R}_I) \langle \mu_n(\mathbf{R}_I) \rangle_{Z(\mathbf{r}_I)}. \tag{112}$$

The ensemble averages implied by the symbol $\langle \ldots \rangle$ are performed by successive electronic structure-based molecular dynamics simulations with specific combinations of the electronic and nuclear chemical potential. In this way, one can explore a path, $\lambda(N_e, Z(\mathbf{r}_I))$, in chemical space passing from one compound to another as if the system was in contact with a reservoir of electrons and protons/nuclei. As we have seen, the GC approach to the electronic systems was developed almost in parallel to the standard DFT with fixed N while for other electronic structure theories GC approaches were developed only more recently [265,266]. The main reason, in our view, is that DFT has a semi-classical character, having $\rho(\mathbf{r})$ as a central quantity, thus analogies with classical statistical mechanics and thermodynamics may be easier. Instead, methods, based on the explicit treatment of the electronic wavefunction, require a full quantum mechanical treatment in terms of creation and annihilation operators in order to add and remove particles. In the next section, we discuss one representative example, that is the treatment of open systems in the Hartree–Fock method [265].

4.4. Variational Grand Canonical procedure for open system Hartree–Fock wavefunctions

While in DFT the analogy with classical thermodynamics is, at least, intuitive the same cannot be said for open boundary systems, where the electronic wavefunction is explicitly taken into account. As we will see, in such approaches, as anticipated above, the change in number of particles must be performed through the explicit use of creation and annihilation operators. Such a procedure implies the re-writing of the Hamiltonian in terms of such operators and this leads to a second-quantized Hamiltonian treatment. In this context, here we report about a variational GC procedure for open system Hartree–Fock wavefunctions. Let us define the nonrelativistic Hamiltonian of electrons for ordinary matter (i.e. in presence of nuclei) in atomic units:

$$H = \sum_n \left[-\frac{\nabla_n^2}{2} + v(\mathbf{r}_n) \right] + \frac{1}{2} \sum_{n \neq m} \frac{1}{r_{nm}} = \hat{T} + \frac{1}{2} \sum_{n \neq m} \frac{1}{r_{nm}}, \tag{113}$$

where, as usual, $v(\mathbf{r}_m)$ is the Coulomb potential, generated by the nuclei and acting on the electrons, and $\frac{1}{2}\sum_{n\neq m}\frac{1}{r_{nm}}$ is the electron–electron Coulomb repulsion. Let us further consider an orthonormal basis of M real single-particle orbitals $\phi(_i(\mathbf{r}))$; $i = 1, \ldots, M$ as usually considered in the Hartree–Fock method (we assume that the reader has knowledge of the

Hartree–Fock method since it is treated in all undergraduate programs of physics and chemistry). The operators, which we consider, are the annihilation and creation operators, such that $\hat{c}_{js}^{+}$ creates an electron of spin s (up or down) in the orbital $\phi_j(\mathbf{r})$, while $\hat{c}_{js}$ annihilates an electron on of spin s (up or down) in the orbital $\phi_j(\mathbf{r})$. The operators have the following anticommutation rules, being electrons fermionic particles of spin $\frac{1}{2}$: $[\hat{c}_{js}^{+}, \hat{c}_{j's'}]_+ = \delta_{ss'}\delta_{jj'}$; $[\hat{c}_{js}, \hat{c}_{j's'}]_+ = [\hat{c}_{js}^{+}, \hat{c}_{j's'}^{+}]_+ = 0$. The Hamiltonian of Eq. (113) can now be written as:

$$\hat{H} = \sum_{ij,s} T_{ij}\hat{c}_{is}^{+}\hat{c}_{js} + \frac{1}{2}\sum_{ijkl,s,s'} V_{ijkl}\hat{c}_{is}^{+}\hat{c}_{ks'}^{+}\hat{c}_{ls'}\hat{c}_{js}, \tag{114}$$

where $T_{ij} = \int \phi_i(\mathbf{r})\hat{T}\phi_j(\mathbf{r})d\mathbf{r}$ (one electron integral) and $V_{ijkl} = \int \phi_i(\mathbf{r})\phi_j(\mathbf{r})v(|\mathbf{r}-\mathbf{r}'|)\phi_i(\mathbf{r}')\phi_j(\mathbf{r}')d\mathbf{r}d\mathbf{r}'$. For simplicity, since T_{ij} and V_{ijkl} do not explicitly refer to the spin, the explicit reference to the spin is thus removed for the formalism in the treatment. Let us introduce the density matrix operator elements:

$$\hat{\rho}_{ij} = \sum_{s=1,2} \hat{c}_{is}^{+}\hat{c}_{js}; \quad \hat{\rho}_{ij}^{+} = \hat{\rho}_{ji}. \tag{115}$$

It follows from Eqs. (114) and (115)

$$\hat{H} = \sum_{ij} t_{ij}\hat{\rho}_{ij} + \frac{1}{2}\sum_{ijkl} \hat{\rho}_{ij}V_{ijkl}\hat{\rho}_{kl}, \tag{116}$$

with $t_{ij} = T_{ij} - \frac{1}{2}\sum_I V_{ilkj}$ where $I = (ij)$ (and $K = (kl)$). At this point, the Gibbs free energy is introduced as:

$$G(\beta, \mu) = -\beta \ln Z(\beta, \mu), \tag{117}$$

where, as usual, $\beta^{-1} = k_B T$, μ is the chemical potential, and $Z(\beta, \mu)$ is the GC partition function:

$$Z(\beta, \mu) = \sum_{N=0}^{2M} e^{\beta\mu N} Tr_N[e^{-\beta\hat{H}}] = Tr[e^{-\beta(\hat{H}-\mu\hat{N})}]. \tag{118}$$

The trace is performed over all antisymmetric N-electron states with all possible numbers of electrons. The maximum number, which can be accommodated on the basis set, is $2M$. $\hat{N}$ is the number operator, whose properties can be found in Appendix G. Due to the electron–electron interaction the calculation of $Z(\beta, \mu)$ is in practice prohibitive, i.e., it requires the sampling on all possible electronic configurations. At this point, Jacobi and Baer propose the innovative aspect of their treatment, that is to shift the problem to noninteracting electrons embedded in a mean-field potential (one-body potential) u. It follows that the noninteracting Hamiltonian can be written as

$$\hat{H}_0 = \sum_{ij} t_{ij} + u_{ij}\hat{\rho}_{ij}. \tag{119}$$

If the matrix $h = t + u$ has eigenvalues $\epsilon_1, \epsilon_2 \ldots$ then we have

$$Z_0(\beta, \mu) = \Pi_{l=1,M}(1 + e^{\beta(\mu-\epsilon_l)})^2, \tag{120}$$

which can be shown to have the form:

$$Z_0(\beta, \mu) = det[1 + e^{\beta(\mu-h)}]^2. \tag{121}$$

While we can now calculate in a simplified way the expression of Eq. (121), the question obviously concerns the realistic character of such an approximation for the true system, i.e., interacting electrons one wants to study. The approach proposed is based on a simple observation:

$$\left\langle \hat{H} - \hat{H}_0 \right\rangle = 0 \rightarrow \frac{1}{2}V_{IK}\left\langle \hat{\rho}_I\hat{\rho}_K \right\rangle - u_I\left\langle \hat{\rho}_I \right\rangle = 0, \tag{122}$$

where $\langle .. \rangle$ denotes the average w.r.t. $\hat{H}_0$. Next, the Gibbs–Peierls–Bogoliubov inequality is used (see, e.g., [59,60] and references therein):

$$G(\beta, \mu) \leq \Gamma(\beta, \mu) \equiv G_0(\beta, \mu) + \left\langle \hat{H} - \hat{H}_0 \right\rangle, \tag{123}$$

where $G_0(\beta, \mu)$ is the free energy of the noninteracting system and the quantity $\Gamma(\beta, \mu)$ is called the effective free energy. By varying h (or better u) in $\hat{H}_0$ one can minimize $\Gamma(\beta, \mu)$ and thus have the best variational approximation of $\Gamma(\beta, \mu)$. Once h (u) is determined, then for any observable one has

$$\left\langle \hat{O} \right\rangle_{exact} \approx \left\langle \hat{O} \right\rangle, \tag{124}$$

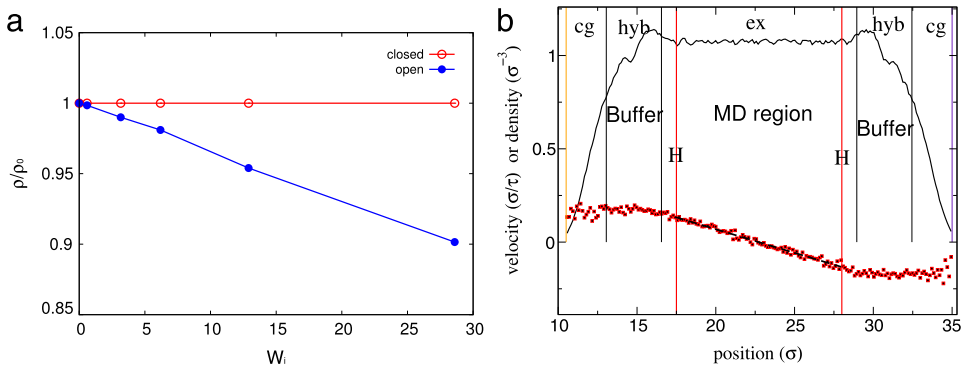

Fig. 18. (a) The density decrease under shear in OBMD simulation. (b) Steady Couette flow of liquid water simulated by the triple-scale scheme coupling atomistic and continuum hydrodynamics.
Source: Reprinted from Refs. [82,225].

that is

$$Tr[e^{-\beta(\hat{H}-\mu\hat{N})}\hat{O}] \approx Tr[e^{-\beta(\hat{H}_0-\mu\hat{N})}\hat{O}]_{u=u_{OPT}}, \tag{125}$$

where $[..]_{u=u_{OPT}}$ indicates that $\hat{H}_0$ is the noninteracting Hamiltonian with optimized mean-field potential u_{OPT}. Eq. (125) allows then for a GC treatment of an electron system described via the Hartree–Fock approach. It has been applied to the study of several systems, for example, molecular hydrogen and water systems [265] and di-lithium systems [267].

5. Open molecular systems out of equilibrium

Recent advances in nanotechnology and nanomedicine have triggered much development in theoretical and simulation approaches to study non-equilibrium systems. The hybrid approaches bridging to continuum hydrodynamics, described in Section 3, are especially useful for simulations of the transport of nanoparticles through fluids, which is a typical example from nanofluidics. Simulations can provide insight into such systems when they can access, both, the atomistic length scales associated with size of the nanoparticles and the micro/macro scales characteristic of the carrier flow field [217]. Simulations using MD can capture the atomistic details of the nanoparticle-liquid interface but due to their computational cost they cannot be extended currently to the macroscale regime of the full flow field. In turn, continuum descriptions, using the Navier–Stokes equations may capture the macro-scale behavior of the fluid flow but they fail to represent accurately the flow field at the nanoparticle surface. The hybrid approaches, on the other hand, combine the powerful features of the both descriptions, i.e., the ability to describe the macro-scale behavior of the flow as well as accurate boundary conditions around nanoparticles [268,269]. Below, we shall give some examples that show the suitability of multiscale methods based on AdResS (The linear momentum preservation and its importance for hydrodynamics are here critical!) to tackle such non-equilibrium scenarios.

5.1. Non-equilibrium OBMD

As already mentioned, OBMD allows us to impose arbitrary time-dependent external pressure tensor. Here, we present a shear flow of the star-polymer melt depicted in Fig. 7 to illustrate the applicability of the method for studying non-equilibrium situations [81,82]. OBMD simulations reveal that shearing polymer melts at constant normal pressure produces different rheology than shearing at a constant volume as shear stress induces significantly different redistribution of pressure in comparison with the closed simulation [82]. This is one of examples, where the new open methodology is actually not expected to agree with a closed simulation using periodic boundary conditions. Instead, it should more faithfully reproduce the experimental setup. Moreover, straining the melt with increasing shear stress also induces melt expansion and consequently density drop, i.e., shear dilatancy, as shown in Fig. 18(a). This phenomenon cannot be studied with the closed simulation using Lees–Edwards boundary conditions [270], as the number of molecules and hence density remain constant [82].

Employing the triple-scale scheme, presented in Section 3.2.3, bridging atomistic and continuum hydrodynamics, one can simulate different types of steady and unsteady fluid flows such as Couette or oscillatory shear flows [219,225]. Fig. 18(b) shows the density and velocity profiles, which agree well with the continuum Navier–Stokes solution (dashed line), obtained at the steady state of the Couette flow [225]. This demonstrates that the hydrodynamics is captured well by the hybrid method.

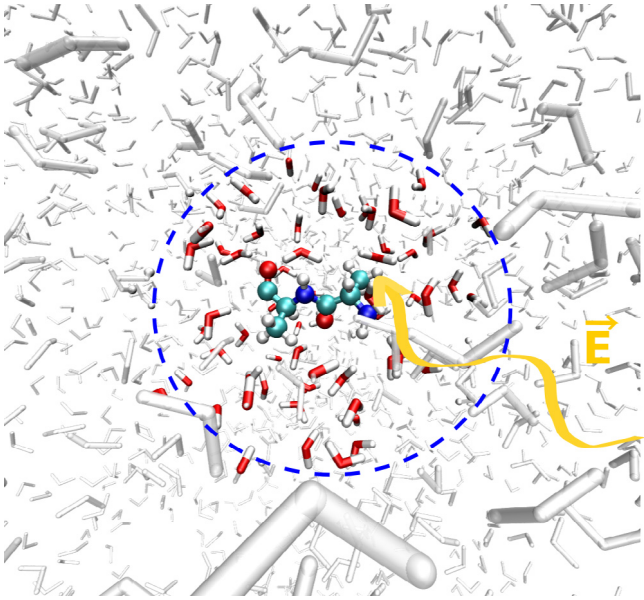

Fig. 19. Pictorial representation of the system investigated. An alanine-dipeptide molecule in solution under the action of an external electric field whose action is localized in the first hydration shells of the molecule. The simulation setup is instead given in Fig. 20.

5.2. Conformational changes of large molecules in solution under the effect of an external perturbation: Towards a non-equilibrium approach in GC-AdResS

While the approach to non-equilibrium, reported in the previous section, deals with collective effects, for other systems of interest, in conditions of non-equilibrium, the focus may be localized in a particular region, e.g., even on a single molecule. For example, a case recently treated is that of the conformational changes of alanine dipeptide in water under the effect of a localized external electric field [271] (see also Fig. 19). This system represents a prototype case for understanding the effects of the electromagnetic radiation on proteins and thus on human tissues. If one considers the solvation region around the solvated molecule, where the electric field acts, as the system of interest and the rest of the system as a reservoir, then the scenario is that typical of a GC setup. The system of interest has a varying number of particles and resembles the idea of GC-AdResS. In the work of Ref. [271], the adaptive character of the approach refers to the fact that molecules entering the region of interest are not subject to the thermostat, whereas molecules outside (i.e. in the reservoir) instead are subject to the action of the reservoir. However, the simulation is carried out at full atomistic level. This is a first step towards the partitioning of space in the AdResS fashion. The reason for defining a subsystem in such a full atomistic system is twofold: (a) As underlined before, the external electric field is localized, thus the analysis is needed only in a subsystem of the whole system. (b) The analysis of the response of the system to the perturbation of the electric field is done in terms of dynamical response and the action of a thermostat on the molecules under observation may introduce artificial contributions to the dynamics. The setup of the simulation box is done according to Fig. 20 and the region of interest is what is referred to as "dynamical region". The simulation methodology employed is the so-called Dynamical Non-Equilibrium Molecular Dynamics (D-NEMD) developed by Ciccotti and coworkers [272–275]. Below, we report the essential features of D-NEMD. For more details, we invite the reader to consult the references given above. Let us denote a macroscopic observable by $O(t)$, at time t. The configurational probability distribution is $\rho(\mathbf{x}, t)$, where $\mathbf{x}$ is the phase space variable. Then we have: $O(t) = \int d\mathbf{x}\, \hat{O}(\mathbf{x})\rho(\mathbf{x}, t) = \langle \hat{O}(\mathbf{x}), \rho(\mathbf{x}, t)\rangle$ with $\hat{O}(\mathbf{x})$ a microscopic observable. It is assumed that the initial probability distribution $\rho(\mathbf{x}, 0)$ is known. In particular, in our case, it is identical to the equilibrium distribution of the system without the external perturbation (i.e. in our case the electric field). The dynamics of the system is governed by the Hamiltonian equation, i.e. $\dot{\mathbf{x}} = \nabla_{\mathbf{x}}\mathcal{H}(\mathbf{x}, t)$, where $\mathcal{H}$ is a time-dependent Hamiltonian. It follows that the Liouville equation for the probability distribution is:

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} = -iL(t)\rho(\mathbf{x}, t), \tag{126}$$

where $iL(t) = \{\cdot, \mathcal{H}\}$ is the Liouville operator. Eq. (126) can be solved in a formal way by $\rho(\mathbf{x}, t) = U^{\dagger}(t, 0)\rho(\mathbf{x}, 0)$, where $U^{\dagger}(t, 0) = \mathcal{T}\exp\{-i\int_0^t dt' L(t')\}$, and $\mathcal{T}$ is the time ordering operator. We have also: $\frac{d\hat{O}(\mathbf{x}(t))}{dt} = \nabla_{\mathbf{x}}\hat{O}\cdot\dot{\mathbf{x}} = \nabla_{\mathbf{x}}\hat{O}\cdot\nabla_{\mathbf{x}}\mathcal{H} =$

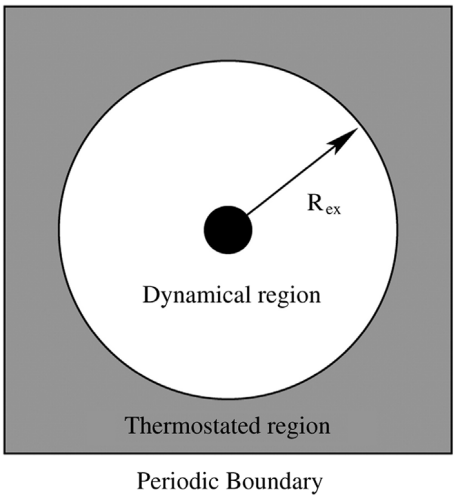

Fig. 20. Pictorial representation of the simulation setup corresponding to Fig. 19. The region of interest with varying number of molecules is named "dynamical region". Molecules in this region are not subject to the action of the thermostat, molecules outside are instead subject to the thermostat. Due to the large reservoir region, the "dynamical region" is properly thermalized.
Source: Reprinted from Ref. [271].

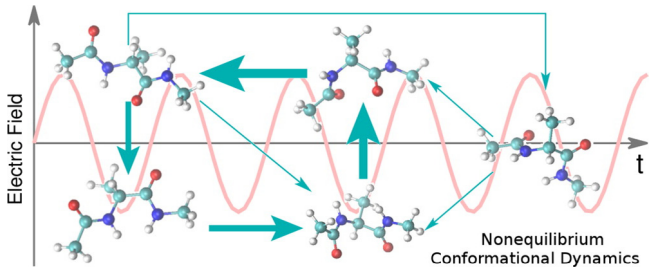

Fig. 21. Schematic representation of the conformational changes of the alanine-dipeptide as a result of the application of an external (in such specific case oscillatory) electric field. The thickness of the arrows is an index of the strength of the flux probability from one conformation to another.
Source: Reprinted from Ref. [271].

$iL(t)\hat{O}(\mathbf{x}(t))$. This equation can be solved by $\hat{O}(\mathbf{x}(t)) = U(t,0)\,\hat{O}(\mathbf{x}(0))$. Hence,

$$O(t) = \langle \hat{O}(\mathbf{x}),\, \rho(\mathbf{x},t)\rangle = \langle \hat{O}(\mathbf{x}),\, U^{\dagger}(t,0)\,\rho(\mathbf{x},0)\rangle = \langle U(t,0)\,\hat{O}(\mathbf{x}),\, \rho(\mathbf{x},0)\rangle$$
$$= \langle \hat{O}(\mathbf{x}(t)),\, \rho(\mathbf{x},0)\rangle. \tag{127}$$

We have assumed that the system starts from the equilibrium distribution (without the effects of the external perturbation). Then, from Eq. (127) one concludes that the observable $O(t)$, calculated under the action of the external field (i.e. in situation of non-equilibrium), is the same as obtained from the ensemble average of the microscopic observable computed along trajectories, starting from initial configurations sampled from an equilibrium distribution. In terms of the numerical algorithm, the procedure consists of first, running an equilibrium MD simulation in order to generate a sample of equilibrium configurations. Next, these configurations are employed as initial configurations for the full dynamics (with the external perturbation acting) and each trajectory is integrated until time t. In simple words, one has a branching of trajectories of non-equilibrium starting from points along the trajectory of equilibrium. The value of the macroscopic observable at time t is obtained by averaging the observable calculated at each time t along each "branching" (non-equilibrium) trajectory. In the application of Ref. [271], constant (instantaneous) and oscillating electric fields are applied and various response functions in the "dynamical region" are analyzed. In this way, one can draw the map of conformational changes of the solvated molecule under the effect of the external field (see e.g. Fig. 21). As anticipated, the next step will be that of defining the reservoir of Fig. 20 as in GC-AdResS.

6. Outlook and perspectives

The studies and approaches, we have reported and discussed in this review, are potentially powerful tools for a decisive step forward in the treatment of classical and quantum systems with open boundaries. Here, we attempt to list some

suggestions of perspectives that come straightforwardly to mind. The projector-operator technique of Section 2.2 (the ES equation), developed within a rigorous quantum formalism, represents a very general theoretical framework within which numerical procedures, for molecular system with open boundaries, can find their formal legitimization. The corresponding (equilibrium) stationary equation, i.e. $\left(\frac{d}{dt} + iL^S_{eff}\right)\rho_S(t) = 0$, involves the Liouvillian operator L_{eff}, in which the interaction between system and reservoir is encoded. This latter is a general term and the system-reservoir coupling operators can be written in terms of creation and annihilation of particles. Such a framework implies that open systems with electrons, for example, can be described rigorously once the process of creation or annihilation is properly defined. In this review, we have presented effective numerical procedures by which the generic formal operators of the theoretical framework above can be implemented in practice (i.e. by properly varying the number of electrons as the system evolves according to its physics), that is through the CNP approach of Section 4.3 and the GC approach of Hartree–Fock in Section 4.4. Moreover, regarding the nuclei, the same equation of the electrons applies, except that, being usually treated as classical objects, the classical limit of the equation can be used. The classical limit of the equation is nothing else than its passage from quantum operators with their complex algebra to standard classical quantities. The BL model of Section 2.1 is nothing else than a classical version of the quantum equation for open systems. We have shown that for classical atoms the BL model represents the theoretical framework necessary for a rigorous definition of physical quantities calculated in simulation, e.g. equilibrium time correlation functions in systems with open boundaries and variable number of molecules. The conclusion is that a general conceptual framework for open boundary systems with molecules described (in the usual way) as classical nuclei and electrons, already exists and is expressed by equation: $\left(\frac{d}{dt} + iL^S_{eff}\right)\rho_S(t) = 0$, where the quantum part (electrons) can be treated via existing methods, which vary the number of electrons and the classical part for the nuclei can be described by the BL method and numerically implemented via, for example, GC-AdResS. Given the difference in time scale, the two equations can be treated in a separate way and adiabatically coupled, as for example done in standard electronic structure calculations. In conclusion, we consider the equation above as a rigorous formal framework for a truly GC treatment of electrons and nuclei in molecular systems and we have linked it to available numerical procedures, which can be combined following its recipe, thus offering a challenging but intriguing research plan. It is also true that its actual implementation in molecular-simulation codes would not be trivial. In any case, the idea of rigorous projection of all variables onto the essential electronic DOFs via a simplified system-reservoir coupling operator is certainly appealing. Although these represent generic suggestions, they nevertheless represent the basis for a substantial research program for years to come. It must also be underlined that in some cases where quantum mechanics is relevant, as for PIMD, the concept of adaptive molecular resolution already offers the possibility of studying open systems at the same conceptual complexity of the classical models. Besides the very general scheme of coupling discussed above, the theoretical section offers other interesting connections to current simulation approaches. For example, for the study of open nanosystems; in fact the problem of the coupling energy between the reservoir and the system can be treated by both the GC-AdResS and OBMD methods since the effective coupling energy can be directly calculated, thus they may be used in the future to access the calculation of properties in small (open) systems as defined by Hill in Section 2.5. Interestingly, the idea of open boundaries makes also possible to consider physical ensembles characterized by intensive quantity only (e.g. μ, p, T). While this is not possible in standard (macroscopic) thermodynamics, it becomes possible in nanothermodynamics. Technical implementations of the adaptive resolution technique, which implicitly use this idea, have already been presented [276]. We can foresee the extension of the current simulation approaches to subjects of highest priority such as the interdisciplinary *stochastic thermodynamics* [277]. There is also a flux of information in the opposite direction, that is the numerical methods inspire theoretical developments. In fact, the GC-AdResS method has in the meanwhile inspired investigation at fundamental level of a relevant problem of statistical mechanics, that is how one can properly partition a large system in non-interacting or weakly interacting systems. This research led to an extension of the rigorous results of the Peierls-Bogoliubov inequality of Section 2.5 to a two sided inequality that bounds from above and below the interacting energy of two subsystems of a large system [61,278], thus lying the basis to go beyond Hill's approach. Finally, the theory of fluctuations of Section 2.6 can be used to improve current numerical techniques, e.g. GC-AdResS, when in presence of a small reservoir. The thermodynamic force, written as a linear function of the gradient of the particle number density and currently assuring the equilibrium density in AdResS simulations, following the principles of Section 2.6, should be extended beyond the linear term and assure the particle number fluctuations as well. Thus, the model for a finite reservoir of Section 2.6 and the model for open nanosystems of Section 2.5 will enlarge the window of scenarios treatable by adaptive molecular dynamics methods and will allow to avoid finite size effects in molecular simulations. Finally, a truly numerical GC approach for the high resolution region can be achieved by either coupling the particle-based approach with the continuous approach, as discussed in Section 3.2.3, or by opening up the simulation box with the OBMD method of Section 3.2.2, where OBMD mimics effectively an infinite reservoir. Methods to study systems with irreversible transport of heat and matter, as that reported in Section 5.1, and in general systems out of equilibrium, will certainly gain in conceptual solidity if framed within the BL or ES models. In principle, from the statistical and dynamical point of view, the ES model, as underlined before, represents a very advanced framework.

7. Conclusions

In this review, we sketched a map of theoretical models and computational methods, available in literature, devised to study molecular systems with open boundaries. The focus is centered around the idea of coupling different resolutions in an open boundary fashion, with a particular emphasis on the advancement of the AdResS approach. Such methods have

meanwhile gone beyond the original purpose of multiscale approaches and are currently employed also for truly Grand Ensemble simulations with ideal or real reservoirs. To complete this aim, the theoretical models, presented in Section 2, are reported as powerful concepts that can complement and integrate the needs of molecular simulation, when dealing with realistic systems. The example of the use of the BL model into the adaptive setup has been discussed in relation to the numerical definition of equilibrium time correlation functions in systems with open boundaries and variable number of molecules. Slightly more delicate is the case of quantum systems. In some cases, as for PIMD, the concept of adaptive molecular resolution already offers the possibility of studying open systems at the same conceptual complexity of the classical models. The problem becomes more challenging when electrons become relevant. While the QM/MM schemes are very practical but conceptually not solid if used to simulate GC systems, methods such as CNP or the GC extension of the Hartree–Fock method provide very solid conceptual ingredients for a truly GC treatment of electrons and nuclei.

In conclusion, although we have reported on a large number of successful applications, ranging from biomolecules in solution to macromolecular liquids of materials science, an optimal embedding of first principle concepts into numerical schemes remains a vivid aim. In this perspective, the aim of this review, enforced by the discussion in the previous section, is simply to offer the ingredients for further development in the field.

Acknowledgments

We would like to thank many collaborators and students who, during the years, have made possible the success of our research concerning open systems. In particular, we would like to thank Kurt Kremer, who encouraged us to start the research in this field, and Rafael Delgado-Buscalioni for many fruitful discussions on topics described in this review. We are also grateful to Luca Ghiringhelli, Christian Gogolin, Ruth M. Lynden-Bell, Julija Zavadlav, Jurij Sablić, and Rudolf Podgornik for critical reading of the manuscript and their constructive remarks. L. D. S. acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 (project C01) and from the European Community through project E-CAM. M. P. acknowledges financial support through the grants P1-0002 and J1-7435 from the Slovenian Research Agency.

Appendix A. Molecular dynamics

In MD simulations [279], we compute the evolution of a system according to the classical Hamilton equations[3]

$$\frac{d\boldsymbol{\eta}}{dt} = \{\boldsymbol{\eta}, H\} = \hat{L}_H \boldsymbol{\eta}, \tag{A.1}$$

where $\widehat{L}_H$ is the Lie operator, $\{,\}$ is the Poisson bracket and $\boldsymbol{\eta} = (\mathbf{q}, \mathbf{p})$ is a vector of the coordinates of all the particles and their conjugate momenta.

The formal solution of the above Hamiltonian system can be written in terms of Lie operators as

$$\boldsymbol{\eta}|_{t+\Delta t} = \exp(\Delta t \hat{L}_H)\boldsymbol{\eta}|_t \tag{A.2}$$

and represents the exact time evolution of a trajectory in phase space composed of coordinates and momenta of all the particles from t to $t + \Delta t$, where Δt is the integration timestep [47].

If we split Hamiltonian H into two terms as $H = U + T$ and use a second order approximation, known as the generalized leap-frog scheme [280,281], for Eq. (A.2)

$$\boldsymbol{\eta}|_{t+\Delta t} = \exp\left(\frac{\Delta t}{2}\hat{L}_U\right)\exp(\Delta t \hat{L}_T)\exp\left(\frac{\Delta t}{2}\hat{L}_U\right)\boldsymbol{\eta}|_t + \mathcal{O}(\Delta t^3), \tag{A.3}$$

we obtain the widely-used velocity Verlet algorithm [47]. If we split H in a different way we obtain different integrators. For example, splitting $H = H_0 + H_r$, where H_0 is the high-frequency harmonic part, which can be solved analytically using the normal modes of a given molecule, and H_r the remainder, yields the Split Integration Symplectic Method (SISM) [127,282–284].

The velocity Verlet algorithm is second-order (only one force calculation per integration timestep) and symplectic, i.e., phase space area conserving (see Liouville's theorem) [47]. Expanding Eq. (A.3) yields the following numerical integration scheme for coordinates and momenta of a given particle i:

$$\mathbf{p}_i\left(t + \frac{1}{2}\Delta t\right) = \mathbf{p}_i(t) + \frac{1}{2}\Delta t \mathbf{F}_i(t), \tag{A.4}$$

$$\mathbf{q}_i(t + \Delta t) = \mathbf{q}_i(t) + \frac{\Delta t}{m}\mathbf{p}_i\left(t + \frac{1}{2}\Delta t\right), \tag{A.5}$$

$$\mathbf{p}_i(t + \Delta t) = \mathbf{p}_i\left(t + \frac{1}{2}\Delta t\right) + \frac{1}{2}\delta t \mathbf{F}_i(t + \Delta t), \tag{A.6}$$

[3] They can also be given in the Lagrange or Newton form (see Section 3.1.2).

where $\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{q}_i}$ is the total force acting on the ith particle. For stochastic MD using, for example, Langevin or DPD thermostats, other integration methods are more appropriate, see, e.g., Ref. [285] and references therein.

MD yields information about temporal evolution of properties of the system. We compute statistical properties of the system as time averages over the trajectories. If we assume that the system is ergodic [47] then these statistical properties match the ones computed from the corresponding microcanonical *NVE* statistical ensemble (computed by, for example, the MC approach) in the thermodynamic limit [286]. To mimic an infinitely large system, MD simulations are traditionally performed under periodic boundary conditions (PBCs) [286]. If a particle leaves the box during the simulation then it enters back at the opposite side of the box. Thus, the total number of particles in the simulation box remains constant. Typically, MD simulations are performed in the microcanonical *NVE* and canonical *NVT* ensembles [68]. However, like the MC technique (see below) it can be modified to sample from other statistical ensembles such as isothermal–isobaric NPT or GC μVT, as explained in Section 3.

Appendix B. Monte Carlo

MC is a stochastic simulation method [66,287,288]. In the Metropolis algorithm, a trial configuration x' of a system is generated by a random perturbation of the initial configuration x (displacement of particles). The trial configuration is accepted if its corresponding energy $U(x') < U(x)$. However, if $\Delta U = U(x') - U(x) > 0$ then the new configuration is accepted with the probability

$$P_{dis} = \min\left(1, \exp(-\Delta U/k_B T)\right), \tag{B.1}$$

that is, if P_{dis} is greater than a uniformly generated random number on interval (0,1). Otherwise, the move is rejected. In this way, we sample configurations with the Boltzmann probability. Ensemble averages are then computed from the obtained set of configurations for properties of the system that depend on positions only. There is no kinetic energy contribution in the total energy, which is determined solely by the potential energy U. There is also no time relationship between successive MC configurations as each new configuration depends only upon the previous configuration [68].

In GC MC simulations, we sample the GC probability distribution. To this end, an additional trial move, i.e., insertion and deletion of particles, is introduced besides the Metropolis displacement move, described above. A new particle is inserted at a random position and the move is accepted with a probability [66]

$$P_{ins} = \min\left(1, \frac{V}{\Lambda^3(N+1)} \exp([\mu - U(N+1) + U(N)]/k_B T)\right), \tag{B.2}$$

where V is the volume, μ is the chemical potential and Λ the thermal de Broglie wavelength. A randomly selected particle is deleted with a probability [66]

$$P_{del} = \min\left(1, \frac{\Lambda^3 N}{V} \exp(-[\mu + U(N-1) - U(N)]/k_B T)\right). \tag{B.3}$$

For further details, see also Section 4.1.1.

Appendix C. Path integral formalism in a nutshell

In this section, we sketch the basic procedure to obtain a quantized Hamiltonian via the path integral formalism within the framework of quantum statistical mechanics. A detailed discussion of both formal and numerical aspects of the theory and of applications can be found in Refs. [47,235] (see also [289]). We start from the Hamiltonian of a single particle subject to an external potential U:

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + U(\hat{x}) = \hat{K} + \hat{U}, \tag{C.1}$$

where $\hat{K}$ and $\hat{U}$ are the kinetic and potential energy operators, respectively, with $[\hat{K}, \hat{U}] \neq 0$. The density matrix element in the space representation is:

$$\rho(x, x') = \langle x'|e^{-\beta\hat{\mathcal{H}}}|x\rangle. \tag{C.2}$$

Since $\hat{K}$ and $\hat{U}$ do not commute, the Trotter theorem is used to facilitate the computation of (C.2). Given two non-commuting operators, A and B, the Trotter theorem states that:

$$e^{\lambda(\hat{A}+\hat{B})} = \lim_{P\to\infty} [e^{\frac{\lambda\hat{A}}{2P}} e^{\frac{\lambda\hat{B}}{P}} e^{\frac{\lambda\hat{A}}{2P}}]^P, \tag{C.3}$$

where P is named "Trotter number". The application of the theorem to our specific case leads to: $\rho(x, x') = \lim_{P\to\infty}\langle x'|[e^{-\frac{\beta\hat{U}}{2P}} e^{-\frac{\beta\hat{K}}{P}} e^{-\frac{\beta\hat{U}}{2P}}]^P|x\rangle$. Let us define the operator $\hat{\Omega} = e^{-\frac{\beta\hat{U}}{2P}} e^{-\frac{\beta\hat{K}}{P}} e^{-\frac{\beta\hat{U}}{2P}}$, it follows:

$$\rho(x, x') = \lim_{P\to\infty} \langle x'|\hat{\Omega}^P|x\rangle = \langle x'|\hat{\Omega}^P|x\rangle = \lim_{P\to\infty} \langle x'|\hat{\Omega}\hat{\Omega}\hat{\Omega}\hat{\Omega}\hat{\Omega}..\hat{\Omega}|x\rangle. \tag{C.4}$$

Next, by $P-1$ insertions between each $\hat{\Omega}$ of the identity operator, $\hat{I}=\int dx|x\rangle\langle x|$, one obtains:

$$\rho(x,x')=\lim_{P\to\infty}\int dx_2..dx_P\times\langle x'|\Omega|x_P\rangle\langle x_P|\Omega|x_{P-1}\rangle\langle x_{P-1}|...|x_2\rangle\langle x_2|\Omega|x\rangle. \tag{C.5}$$

The problem is now reduced to the calculation of the matrix elements $\langle x_{k+1}|\hat{\Omega}|x_k\rangle = \langle x_{k+1}|e^{-\frac{\beta\hat{U}}{2P}}e^{-\frac{\beta\hat{K}}{P}}e^{-\frac{\beta\hat{U}}{2P}}|x_k\rangle$. Since $\hat{U}|x_k\rangle = U(x_k)$, one obtains: $\langle x_{k+1}|e^{-\frac{\beta\hat{U}}{2P}}e^{-\frac{\beta\hat{K}}{P}}e^{-\frac{\beta\hat{U}}{2P}}|x_k\rangle = e^{-\frac{\beta U(x_{k+1})}{2P}}\langle x_{k+1}|e^{-\frac{\beta\hat{K}}{P}}|x_k\rangle e^{-\frac{\beta U(x_k)}{2P}}$. Regarding the kinetic operator one can introduce the identity operator in momentum space $I=\int dp|p\rangle\langle p|$ and, since $\hat{K}|p\rangle=\frac{p^2}{2m}$, obtain: $\langle x_{k+1}\mid e^{-\frac{\beta\hat{K}}{P}}\mid x_k\rangle=\int e^{-\frac{\beta p^2}{2mP}}\langle x_{k+1}\mid p\rangle\langle p\mid x_k\rangle$. Next, using the relation between the position and momentum eigenstate: $\langle x\mid p\rangle=\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{ipx}{\hbar}}$ one obtains:

$$\langle x_{k+1}\mid e^{-\frac{\beta\hat{K}}{P}}\mid x_k\rangle=\frac{1}{2\pi\hbar}\int dp e^{\frac{ip(x_{k+1}-x_k)}{\hbar}}e^{-\frac{\beta p^2}{2mP}}, \tag{C.6}$$

whose integration delivers the following expression: $\langle x_{k+1}\mid e^{-\frac{\beta\hat{K}}{P}}\mid x_k\rangle=\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{1/2}e^{-\frac{mP}{2\beta\hbar^2}(x_{k+1}-x_k)^2}$ and, by substituting in the expression of the density matrix, one obtains:

$$\begin{aligned}\rho(x,x') = &\lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar}\right)\int dx_2...dx_P\\ &\times\exp\left(-\frac{1}{\hbar}\sum_{k=1}^{P}\left[\frac{mP}{2\beta\hbar^2}(x_{k+1}-x_k)^2+\frac{\beta\hbar}{2P}(U(x_{k+1})+U(x_k))\right]\right)\Bigg|_{x_1=x}^{x_{P+1}=x'}.\end{aligned} \tag{C.7}$$

The partition function corresponds to the trace of the density matrix:

$$Z=\int_0^L dx\langle x|e^{-\beta\hat{\mathcal{H}}}|x\rangle=\int_0^L dx\rho(x,x) \tag{C.8}$$

within an interval $[0,L]$. The explicit expression is:

$$Z=\lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int_{D(L)}dx_1...dx_P\times\exp\left(-\frac{1}{\hbar}\sum_{k=1}^{P}\left[-\frac{mP}{2\beta\hbar}(x_{k+1}-x_k)^2+\frac{\beta\hbar}{P}U(x_k)\right]\right)\Bigg|_{x_1=x}^{x_{P+1}=x'}, \tag{C.9}$$

where $D(L)$ are all paths within $[0,L]$. If we express it as: $Z=\lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int_{D(L)}dx_1...dx_Pe^{-\beta H(x_1...x_P)}$ one can notice that the "effective" Hamiltonian is given by: $H(x_1...x_P)=\sum_{k=1}^{P}\left[\frac{1}{2}m\omega_P^2(x_{k+1}-x_k)^2+\frac{1}{P}U(x_k)\right]$, with $\omega_P=\frac{\sqrt{P}}{\beta\hbar}$ and $x_{P+1}=x_1$. The expression above is known as the discretized path integral (quantum) representation of the partition function of a single particle. In effective terms it represents the partition function of a polymer ring with P beads harmonically linked, with coupling strength ω_P, via nearest neighbor connections. The extension to a system of N particles is not trivial since the particle's statistics/symmetry (fermionic or bosonic) needs to be included. However, for the cases treated in this work, the approach is semiclassical, so in first approximation there is not the necessity of introducing the symmetry. It follows that for a Hamiltonian of N particles in d-dimensions: $\hat{\mathcal{H}}=\sum_{i=1}^{N}\frac{\hat{p}_i^2}{2m}+U(\hat{\mathbf{r}}_1,\ldots,\hat{\mathbf{r}}_N)$, one has:

$$\begin{aligned}Z = &\lim_{P\to\infty}\prod_{i=1}^{N}\left(\frac{m_iP}{2\pi\beta\hbar^2}\right)^{dP/2}\int\prod_{i=1}^{N}dr_i^{(1)}...dr_i^{(P)}\\ &\times\exp\left(-\sum_{k=1}^{P}\left[\sum_{i=1}^{N}\frac{m_iP}{2\beta\hbar^2}(r_i^{(k+1)}-r_i^{(k)})^2+\frac{\beta}{P}U(\mathbf{r}_1^{(k)},\ldots,\mathbf{r}_N^{(k)})\right]\right)_{\mathbf{r}_i^{(P+1)}=\mathbf{r}_i^{(1)}}.\end{aligned} \tag{C.10}$$

A consequence of the formalism is that the beads with the same index k do interact with each other while the cross-interactions are not included. The partition function obtained above can be sampled using MC or MD methods. While, given the Hamiltonian, the MC technique may be applied straightforwardly, for MD one needs to make a further step of derivation.

Appendix D. Path integral formalism in molecular dynamics

In order to perform MD simulations, it is required to have momenta associated to each particle (atom). This aim can be achieved by writing: $\left(\frac{mP}{2\pi\beta\hbar^2}\right)=\int dp_1...dp_P\left(-\beta\sum_{i=1}^{P}\frac{p_i^2}{2m'}\right)$ that is by adding P-Gaussian integrals via fictitious momentum variables $p_1,\ldots,p_P$ with $m'=\frac{mP}{(2\pi\hbar)^2}$, being an arbitrary mass parameter. For simplicity, we consider here a system composed of equivalent atoms, thus $m_j=m_a$ for the j-atom and as a consequence $m_j'=m'$ for the fictitious mass. The extension to

different species of atoms is straightforward. In the text, we will implicitly assume it. With this manipulation one obtains the following partition function:

$$Z = \lim_{P\to\infty} \int dp_1 ... dp_P \int dx_1 ... dx_P \times \exp\left(-\beta \sum_{k=1}^{P} \left[\frac{p_k^2}{2m'} + \frac{1}{2} m_a \omega_P^2 (x_{k+1} - x_k)^2 + \frac{\beta\hbar}{P} U(x_k)\right]\right)\Bigg|_{x_1=x}^{x_{P+1}=x'}, \tag{D.1}$$

where the corresponding quantized Hamiltonian is : $\mathcal{H}_P = \sum_{k=1}^{P}\left[\frac{[\mathbf{p}^{(k)}]^2}{2m'} + \frac{1}{2} m_a \omega_P^2 (x_{k+1} - x_k)^2 + \frac{1}{P} U(x_k)\right]$. According to the expression above, one can perform MD of a system of ring polymers whose fictitious dynamics (due to the introduction of fictitious momenta) allows for a sampling of a quantum partition function. For dynamical properties, the direct Path Integral MD cannot be used. Specific techniques which employ as basis the Path Integral MD were developed to calculate also dynamical properties. Two techniques must be mentioned: (a) Ring Polymer molecular dynamics (RPMD) [290,291] and (b) Centroid Molecular Dynamics (CMD) [292,293].

D.0.1. Ring polymer molecular dynamics

The method was developed by Craig and Manolopoulos [290]. The essential point is that the (actual) physical mass of the atom is used instead of a fictitious mass, that is: $m' = m_a$, we refer to such method in the text as $\mathbf{H}_2$. There exists also an alternative formulation of RMPD [294] where the Hamiltonian considered is:

$$\mathcal{H}_P = \sum_{i=1}^{P}\left[\sum_{j=1}^{N} \frac{[\mathbf{p}^{(i)}]_j^2}{2m_j} + \sum_{j=1}^{N} \frac{m_j}{2\beta_P^2\hbar^2}(\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + U(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i)\right], \tag{D.2}$$

where $\beta_P = \beta/P$. This implies that the actual simulation is performed at P times the original temperature while the harmonic bead–bead interaction and the potential energy terms are no more scaled by P. We refer to this method in the text as $\mathbf{H}_3$. It can be shown that RPMD provides an accurate approximation to the Kubo-transformed correlation functions between operators $\hat{A}$ and $\hat{B}$ [295,296]: $K_{AB}(t) = \frac{1}{\beta Z}\int_0^\beta d\lambda \left[e^{-(\beta-\lambda)\hat{\mathcal{H}}}\hat{A}e^{-\lambda\hat{\mathcal{H}}}e^{i\hat{\mathcal{H}}t/\hbar}\hat{B}e^{-i\hat{\mathcal{H}}t/\hbar}\right]$. The RPMD approximation in this case is given by [291]:

$$\tilde{c}_{AB}(t) \approx \frac{1}{(2\pi\hbar)^{9PN} Z_P}\int\int d^P p_0 d^P r_0 e^{-\beta_P \mathcal{H}_P(p_0, r_0)} \frac{1}{N}\sum_{i=1}^{N} A_P^i(r_0) B_P^i(r_t), \tag{D.3}$$

where Z_P is the canonical partition function and r_t indicates the position at time t. $A_P(r)$ and $B_P(r)$ are calculated as: $A_P(r) = \frac{1}{P}\sum_{j=1}^{P} A(r_j)$: $\quad B_P(r) = \frac{1}{P}\sum_{j=1}^{P} B(r_j)$. The beads in the polymer ring are treated as dynamical variables and thus are not thermostated (NVE simulations).

D.0.2. Centroid Molecular Dynamics

Centroid Molecular Dynamics (CMD) [292] allows for a reasonable approximation of real time quantum dynamics. A centroid corresponds to a semi-classical object defined as an average over all the beads in a polymer ring: $x_c = \frac{1}{P}\sum_{i=1}^{P} x_i; \quad p_c = \frac{1}{P}\sum_{i=1}^{P} p_i$. The time evolution of x_c occurs according to the equations:

$$\dot{x}_c = \frac{p_c}{m}, \tag{D.4}$$

$$m_c \ddot{x}_c = -\frac{\partial V_o(x_c)}{\partial x_c}, \tag{D.5}$$

where m_c is the physical mass while V_o is a mean field potential generated by the dynamics of the beads. Simulation is made possible by performing the dynamics in normal mode coordinates and by adiabatic decoupling of the fictitious motion of the non-centroid modes from the physical motion of the centroid [297–299]. The quantum time correlation between two operators $\hat{A}$ and $\hat{B}$, via Kubo transform is approximated by [299]:

$$C_{AB}(t) = \frac{1}{Z}\int \frac{dx_c dp_c}{2\pi\hbar} A(x_c(0)) B(x_c(t)) e^{-\beta\mathcal{H}_c}, \tag{D.6}$$

with $\mathcal{H}_c = p_c^2/2m + V_o(x_c)$.

Appendix E. Basics of QM/MM

QM/MM was born as a multiscale approach where only one portion of a system is treated at quantum mechanical level (QM) while the rest of the system is represented at classical level (molecular mechanics, i.e., MM). MM is computationally far less expensive than quantum chemical calculations, either with typical wavefunction based quantum chemistry

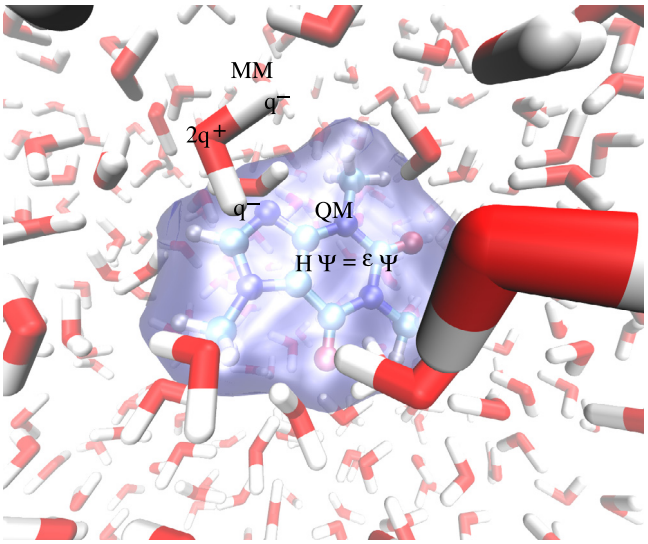

Fig. E.22. Pictorial representation of a typical QM/MM setup. A caffeine molecule in aqueous solution. The region around the solute is treated at quantum level and the electronic charge is continuously distributed in space as a result of the quantum calculation. Outside the QM region water molecules are treated at classical level, here represented as rigid neutral molecules with effective charges $2q^+$, q^-, q^- on the atomic sites.

approach [300] or with Density Functional/Kohn–Sham approaches [255,256] (see also Fig. E.22 as an illustrative example). The physical justification of the idea is based on the fact that electronic properties are very often localized in space and need only a small portion of the system [301]. The fact that the cost of the calculations is drastically reduced compared to a full QM calculation and that results were satisfactory made such methods very popular [246,247]. From the technical point of view, standard QM/MM methods employ a so-called subtractive or additive scheme. For the purpose of this review, in the context of open boundary techniques and adaptive resolution QM/MM, only the additive scheme is of importance and will be described below. The system is described by a global hybrid Hamiltonian: $H_{glob} = H_{QM} + H_{MM} + H_{QM/MM}$ and following this definition the energy of the system (in the ground state) corresponds to the lowest eigenvalue of H_{glob}, so that the QM calculation is done with the presence of the MM environment. The electrostatic coupling of the classical molecules with the quantum molecules is usually taken into account either by classical point charge interactions (i.e., the electronic charges in the QM region are localized on some sites), or by considering the electrostatic interaction of the classical charges as an external potential which polarizes the electronic charge of the QM system. Several techniques are then used to consider bonds occurring across the QM/MM boundaries (see e.g. [302,303]). The major problem with such an approach is the abrupt change in resolution across the QM/MM interface. This is a problem not only for pragmatic questions (bonds that cross the boundaries) but also for conceptual/numerical questions linked to the QM finite size effects. For example, molecules in the QM and molecules in the MM region are not allowed to cross the boundary and change resolution. If the QM region is large enough the problem may be, at least numerically, not important, but usually the QM region is relatively small and the fluctuation of number of molecules (e.g., of the solvent) will have relevant statistical and thermodynamical consequences, altering in turn the electronic properties. The problem of suppression of particle number fluctuation was the main driving force for the development of adaptive resolution QM/MM which is discussed in 4.2 and is of interest to us as basic prototype of open boundary approach for electrons and nuclei.

Appendix F. Kohn–Sham density functional theory: essentials

The Hohenberg-Kohn (HK) formulation of DFT [255,256] has provided the essential platform to access many-electron problems otherwise intractable with other many-electron approaches. The essence of the HK formulation of DFT is the shift from the $3N$-dimensional electronic wavefunction of N electrons, $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$, to the 3-dimensional electron density, $\rho(\mathbf{r}) = \int_{\Omega_{N-1}} |\psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)|^2 d\mathbf{r}_2 \ldots d\mathbf{r}_N$; where Ω_{N-1} indicates that the integral is made over all the domains except that of one particle. This shift, expressed rigorously by the two HK theorems [257], leads to the existence of a variational problem: $E_0 = Min_\rho E[\rho]$; where E_0 is the ground state energy of a system with fixed number of electrons N, and $E[\rho] = \mathcal{F}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})$ is the energy functional of the electron density, with $v(\mathbf{r})$ being the external potential (e.g. nuclei–electron Coulomb interaction). The expression of $E[\rho]$ leads in turn to the existence of a universal functional $\mathcal{F}[\rho] = T[\rho] + V_{ee}[\rho]$ common to all N-electron systems regardless of the external potential. $\mathcal{F}[\rho]$ is composed by two terms, the kinetic functional $T[\rho]$ and the electron–electron Coulomb functional $V_{ee}[\rho]$. Unfortunately, the form of $T[\rho]$ and $V_{ee}[\rho]$ is unknown except, in good approximation, in simplified physical situations. A solution to this problem was later on provided by Kohn and Sham and nowadays known as the Kohn–Sham (KS) approach [258]. They introduce $\frac{N}{2}$ single particle orbitals in a non

interacting frame, $\phi_i(\mathbf{r})$, each accommodating two electrons according to the Pauli principles; in this case one also has: $\rho(\mathbf{r}) = \sum_{i=1}^{\frac{N}{2}} |\phi_i(\mathbf{r})|^2$. This leads to the simplification of the kinetic functional $T[\rho]$ which now can be written exactly in its non interacting form $T_s[\phi] = \sum_{i=1}^{\frac{N}{2}} |\nabla|\phi_i(\mathbf{r})|^2$. Also, $V_{ee}[\rho]$ is simplified and reduced to the Hartree term, $V_{ee}[\rho] = \int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'$ Next, the unknown part of $\mathcal{F}[\rho]$, due to the missing interacting part of the orbitals, is contained in the so called exchange and correlation energy term, $E_{xc}[\rho]$. This setting reduces the HK variational problem to a system of $\frac{N}{2}$ Schrödinger-like equation: $\frac{\hbar^2}{2m}\nabla^2\phi_i(\mathbf{r}) + v_{eff}(\rho, \mathbf{r})\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})$. Here, $\hbar$ is the Planck's constant, m the electron mass, ϵ_i the equivalent of an eigenvalue for the ith orbital. Finally, $v_{eff}(\rho, \mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{exc}(\rho, \mathbf{r})$, with $v_{exc}(\rho, \mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta\rho}$. The advantage of such a formulation is that it provides a single electron description which is rather close to the idea of electron orbitals of chemistry and band-structure idea of solid state physicists. For these reasons, the KS approach gained enormous popularity in the last decades and is today the most used approach in electronic structure calculations. However, a clear disadvantage of the KS approach is implicit in the very formulation of the method: the problem of the unknown $\mathcal{F}[\rho]$ in the HK formulation is now shifted to the problem of the unknown $E_{xc}[\rho]$. This latter can be build, at various levels of approximation, from combining physical intuition, mathematical prescription and numerical data from high level quantum chemical calculations (see e.g. [304–307]). The extension to finite temperature (and to the Grand Potential functional of a Grand Canonical Ensemble) was done by Mermin [256,259] and implies the passage from $\mathcal{F}[\rho]$ to $\mathcal{F}_{gc}[\rho]$. The latter is formally defined as:

$$\mathcal{F}_{gc}[\rho] = Min_{\Gamma\to\rho} Tr\left[\hat{\Gamma}\left(\hat{T} + \hat{V_{ee}} + \frac{1}{\beta}\ln\hat{\Gamma}\right)\right], \tag{F.1}$$

where $\hat{\Gamma} = \sum_N\sum_i |\Psi_{Ni}\rangle\langle\Psi_{Ni}|$, is the density operator in Fock space, i is the index for a particular state with N_i electrons (occupation), N is the total number of electrons, $|\Psi_i\rangle$ is a basis set in Fock space (see also Appendix G), $\hat{T}$ is the standard kinetic operator and $\hat{V_{ee}}$ is the electron–electron Coulomb operator (see e.g. Ref. [308] for a practical use of (F.1)). Instead, an approach to treat fractional number of electrons (for open systems) in DFT is that of considering Kohn–Sham equations with orbital occupation numbers $0 \le f_i \le 1$ and including it into the energy expression through the kinetic energy and the charge density (see [256], see also [309–311]). In this approach, we have:

$$E[\rho, \phi_i] = -\frac{1}{2}\sum_i f_i\langle\phi_i|\nabla^2|\phi_i\rangle + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + V_{ee}[\rho], \tag{F.2}$$

where we have: $\rho(\mathbf{r}) = \sum_i f_i|\phi_i|^2$. The Kohn–Sham setup consists no more of a single determinant but of linear combination of pure KS Slater determinants. The minimization of the energy w.r.t. the occupation number is justified by Janak's theorem [312]:

$$\frac{\partial E[\rho]}{\partial f_i} = \epsilon_i, \tag{F.3}$$

with ϵ_i being the energy of the ith orbital with occupation number f_i. Since the occupation numbers can be fractional, the minimization of the energy of the system can be achieved by transferring an infinitesimal amount of charge (electron) from high energy orbitals to low energy orbital. Following Schneider and Auer [263], the essential distinction between the canonical and GC ensembles is that in the former, one assumes a fixed (constant) number of electrons for each microscopic configuration that the system takes; in such a case the chemical potential is an average over all the configurations. Instead, in the GC ensemble the chemical potential is fixed for each microscopic configuration. Thus, the number of electrons is variable and one can talk only of average number of electrons per configuration. In the work of Schneider and Auer, one also finds technical details (and corresponding references) of further computational approaches that directly employ the concepts above.

Appendix G. Number operator

In general, the number operator is defined as:

$$\hat{N} = \sum_k \hat{N}_k \equiv \sum_k a^+(\phi_k)a(\phi_k), \tag{G.1}$$

that is the product of the creation and annihilation operator acting on a Fock state $|\psi\rangle$ in a basis set $|\phi_k\rangle$: $|\psi\rangle = |\phi_1, \phi_2, \ldots, \phi_m\rangle$; N_k is the number of particles in the state $|\phi_k\rangle$. The action of the number operator follows the standard rules of quantization of the creation and annihilation operator: $\hat{N}_k|\psi\rangle = \sqrt{N_k}a_k^+|\phi_1, \phi_2, \ldots, \phi_{k-1}, \phi_{k+1}, \ldots\ldots\ldots, \phi_m\rangle = \sqrt{N_k}\sqrt{N_k}|\phi_1, \phi_2, \ldots, \phi_{k-1}, \phi_k, \phi_{k+1}, \ldots\ldots\ldots, \phi_m\rangle$, see Refs. [289,313,314] for more details.

References

[1] H. Breuer, F. Petruccione, The Theory of Open Quantum Systems, Oxford University Press, 2002.
[2] J. Gemmer, M. Michel, G. Mahler, Quantum Thermodynamics, in: Lecture Notes in Physics, vol. 784, Springer, 2009.

[3] J. Eisert, M. Friesdorf, C. Gogolin, Quantum many-body systems out of equilibrium, Nat. Phys. 11 (2015) 124.
[4] A. Rivas, S. Huelga, Open Quantum System: An Introduction, in: Springer Briefs in Physics, Springer, Berlin, 2012.
[5] L. Stella, C. Lorenz, L. Kantorovic, The generalized Langevin equation: An efficient approach to non-equilibrium molecular dynamics of open systems, Phys. Rev. B 89 (2014) 134303.
[6] C. Schütte, M. Sarich, Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches, in: Courant Lecture Notes, vol. 89, AMS and the Courant Institute of Mathematical Sciences, 2013.
[7] R. Kapral, Quantum dynamics in open quantum-classical systems, J. Phys.: Condens. Matter 27 (2015) 073201.
[8] H. Quian, Phosphorylation energy hypothesis: Open chemical systems and their biological functions, Annu. Rev. Phys. Chem. 58 (2007) 113.
[9] W. Heuett, H. Qian, Grand canonical markov model: A stochastic theory for open nonequilibrium biochemical networks, J. Chem. Phys. 124 (2006) 044110.
[10] J. Bordin, A. Diehl, M. Barbosa, Y. Levin, Ion fluxes through nanopores and transmembrane channels, Phys. Rev. E 85 (2012) 031914.
[11] S. Chempath, L. Clark, R. Snurr, Two general methods for grand canonical ensemble simulation of molecules with internal flexibility, J. Chem. Phys. 118 (2003) 7635.
[12] S. Boinepalli, P. Attard, Grand canonical molecular dynamics, J. Chem. Phys. 119 (2003) 12769.
[13] N.B. Wilding, P. Sollich, Grand canonical molecular dynamics, J. Chem. Phys. 116 (2002) 7116.
[14] N. Duff, B. Peters, Nucleation in a potts lattice gas model of crystallization from solution, J. Chem. Phys. 131 (2009) 184101.
[15] T. Hill, R. Chamberlin, Extension of the thermodynamics of small systems to open metastable staes: An example, Proc. Natl. Acad. Sci. 95 (1998) 12779.
[16] U. Lucia, A link between nano- and classical thermodynamics: Dissipation analysis (the entropy generation approach in nano-thermodynamics), Entropy 17 (2015) 1309.
[17] N. Halpern, J. Renes, Beyond heat baths: Generalized resources theories for small-scale thermodynamics, Phys. Rev. E 93 (2016) 022126.
[18] Y. Leng, Y. Xiang, Y. Lei, Q. Rao, A comparative study by the grand canonical Monte Carlo and molecular dynamics simulations on the squeezing behavior of nanometers confined liquid films, J. Chem. Phys. 139 (2013) 074704.
[19] T. Prosen, M. Žnidarič, Eigenvalue statistics as an indicator of integrability of nonequilibrium density operators, Phys. Rev. Lett. 111 (2013) 124101.
[20] B. Buča, T. Prosen, Exactly solvable counting statistics in open weakly coupled interacting spin systems, Phys. Rev. Lett. 112 (2014) 067201.
[21] T. Prosen, Exact nonequilibrium steady state of an open hubbard chain, Phys. Rev. Lett. 112 (2014) 030603.
[22] T. Prosen, Quasilocal conservation laws in {XXZ} spin-1/2 chains: Open, periodic and twisted boundary conditions, Nuclear Phys. B 886 (2014) 1177–1198.
[23] W. Frensley, Boundary conditions for open quantum systems driven far from equilibrium, Rev. Modern Phys. 62 (1990) 745.
[24] A. Abramo, Modeling electron transport in mosfet devices: Evolution and state of the art, in: T. Grasser (Ed.), in: In Advanced Device Modeling and Simulation, World Scientific, 2003, pp. 1–27.
[25] I. Tavernelli, R. Vuilleumier, M. Sprik, Ab initio molecular dynamics for molecules with variable numbers of electrons, Phys. Rev. Lett. 88 (2002) 213002.
[26] O. von Lilienfeld, M. Tuckerman, Molecular grand-canonical ensemble density functional theory and exploration of chemical space, J. Chem. Phys. 125 (2006) 154104.
[27] O. von Lilienfeld, Towards the computational design of compounds from first principles, in: V. Bach, L. Delle Site (Eds.), Many-Electron Approaches in Physics, Chemistry and Mathematics, Springer Verlag, 2014, pp. 169–189.
[28] D. Politzer, Condensate fluctuations of a trapped, ideal bose gas, Phys. Rev. A 54 (1996) 5048.
[29] C. Herzog, M. Olshanii, Trapped bose gas: The canonical versus grand canonical statistics, Phys. Rev. A 55 (1997) 3254.
[30] N. Balazs, T. Bergeman, Statistical mechanics of ideal bose atoms in a harmonic trap, Phys. Rev. A 58 (1998) 2359.
[31] A. Daley, Quantum trajectories and open many-body quantum systems, Adv. Phys. 63 (2014) 77.
[32] B. Misra, E. Sudarshan, The zenos paradox in quantum theory, J. Math. Phys. 18 (1977) 756.
[33] N. Syassen, D. Bauer, M. Lettner, T. Volz, D. Dietze, J. Garcia-Ripoll, J. Cirac, G. Rempe, S. Dürr, Quantum trajectories and open many-body quantum systems, Science 320 (2008) 1329.
[34] T. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, J. O'Brien, Quantum computers, Nature 464 (2010) 45.
[35] J. Lebowitz, P. Bergmann, Irreversible gibbsian ensembles, Ann. Phys. 1 (1957) 1.
[36] G. Emch, G. Sewell, Nonequilibrium statistical mechanics of open systems, J. Math. Phys. 9 (1968) 946.
[37] P. Bergmann, J. Lebowitz, New approach to nonequilibrium processes, Phys. Rev. 99 (1955) 578.
[38] J. Lebowitz, A. Shimony, Statistical mechanics of open systems, Phys. Rev. 128 (1962) 1945.
[39] U. Fano, Description of states in quantum mechanics by density matrix and operator techniques, Rev. Modern Phys. 29 (1957) 74.
[40] R. Zwanzig, On the identity of three generalized master equations, Physica 30 (1964) 1109.
[41] J. Seke, Equations of motion in nonequilibrium statistical mechanics of open systems, Phys. Rev. A 21 (1980) 2156.
[42] B. Robertson, Equations of motion in nonequilibrium statistical mechanics, Phys. Rev. 144 (1966) 151.
[43] B. Robertson, Equations of motion in nonequilibrium statistical mechanics. II. Energy Transport, Phys. Rev. 160 (1967) 175.
[44] D. Zubarev, Nichtgleichgewichts-statistische operatoren und quasimittelung in der theorie irreversibler prozesee, Fortschr. Phys. 20 (1972) 471.
[45] D. Zubarev, Grenzbedingungen für statistische Operatoren in der Theorie der Nichtgleichgewichtsprozesse und das Quasimittel, Fortschr. Phys. 20 (1972) 485.
[46] L. Delle Site, Formulation of Liouville's theorem for grand ensemble molecular simulations, Phys. Rev. E 93 (2016) 022130.
[47] M.E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, Oxford University Press, New York, 2010.
[48] M. Peters, An extended Liouville equation for variable particle number systems,1998. arXiv:physics/9809039.
[49] G. Lindblad, On the generators of quantum dynamical semigroups, Comm. Math. Phys. 48 (1976) 119.
[50] A. Kossakowski, On quantum statistical mechanics of non-Hamiltonian systems, Rep. Math. Phys. 3 (1972) 274.
[51] T. Hill, Thermodynamics of small systems, J. Chem. Phys. 36 (1962) 3182.
[52] T. Hill, Thermodynamics of Small Systems, Parts I & II, Dover, 1964.
[53] T. Hill, Perspective: nanothermodynamics, Nano Lett. 1 (2001) 111.
[54] T. Hill, A different approach to nanothermodynamics, Nano Lett. 1 (2001) 273.
[55] S. Schnell, T. Vlugt, J.-M. Simon, D. Bedeaux, S. Kjelstrup, Thermodynamics of a small system in a μT reservoir, Chem. Phys. Lett. 504 (2011) 199.
[56] S. Schnell, X. Liu, J.-M. Simon, A. Bardow, D. Bedeaux, S. Kjelstrup, Calculating thermodynamic properties from fluctuations at small scales, J. Phys. Chem. B 115 (2011) 10911.
[57] S. Schnell, T. Vlugt, J.-M. Simon, D. Bedeaux, S. Kjelstrup, Thermodynamics of small systems embedded in a reservoir: a detailed analysis of finite size effects, Mol. Phys. 110 (2012) 1069.
[58] K. Huang, Statistical Mechanics, Wiley & Son, 1987.
[59] J. Lebowitz, E. Lieb, Existence of thermodynamics for real matter with coulomb forces, Phys. Rev. Lett. 22 (1969) 631.
[60] J. Lebowitz, E. Lieb, The constitution of matter: Existence of thermodynamics for systems composed of electrons and nuclei, Adv. Math. 9 (1972) 316.

[61] L. Delle Site, G. Ciccotti, C. Hartmann, Partitioning a macroscopic system into independent subsystems, J. Stat. Mech. Theory Exp. (2017) in press. http://arxiv.org/abs/1703.10890.
[62] Y. Mishin, Thermodynamic theory of equilibrium fluctuations, Ann. Phys. 363 (2015) 48.
[63] Y. Mishin, Calculation of the γ/γ' interface free energy in the Ni-Al system by the capillary fluctuation method, Model. Simul. Mater. Sci. Eng. 363 (2015) 48.
[64] B. Sadigh, P. Erhart, A. Stukowski, A. Caro, E. Martinez, L. Zepeda-Ruiz, Scalable parallel Monte Carlo algorithm for atomistic simulations of precipitation in alloys, Phys. Rev. B 85 (2012) 184203.
[65] B. Sadigh, P. Erhart, Calculation of excess free energies of precipitates via direct thermodynamic integration across phase boundaries, Phys. Rev. B 86 (2012) 134204.
[66] D. Frenkel, B. Smit, Understanding Molecular Simulation, second ed., Academic Press, San Diego, 2002.
[67] B. Widom, Some topics in the theory of fluids, J. Chem. Phys. 39 (1963) 2808–2812.
[68] A.R. Leach, Molecular Modeling; Principles and Applications, Springer, New York, 2010.
[69] T. Cagin, B.M. Pettitt, Grand molecular dynamics: a method for open systems, Mol. Simul. 6 (1991) 5–26.
[70] T. Cagin, B.M. Pettitt, Molecular dynamics with a variable number of molecules, Mol. Phys. 72 (1991) 169–175.
[71] J. Ji, T. Cagin, B.M. Pettitt, Dynamic simulations of water at constant chemical potential, J. Chem. Phys. 96 (1992) 1333–1342.
[72] S. Weerasinghe, B.M. Pettitt, Ideal chemical potential contribution in molecular dynamics simulations of the grand canonical ensemble, Mol. Phys. 82 (1994) 897–912.
[73] G.C. Lynch, B.M. Pettitt, Grand canonical ensemble molecular dynamics simulations: Reformulation of extended system dynamics approaches, J. Chem. Phys. 107 (1997) 8594–8610.
[74] B.J. Palmer, C. Lo, Molecular dynamics implementation of the gibbs ensemble calculation, J. Chem. Phys. 101 (1994) 10899–10907.
[75] C. Lo, B. Palmer, Alternative hamiltonian for molecular dynamics simulations in the grand canonical ensemble, J. Chem. Phys. 102 (1995) 925–931.
[76] R.M. Shroll, D.E. Smith, Molecular dynamics simulations in the grand canonical ensemble: Application to clay mineral swelling, J. Chem. Phys. 111 (19) (1999) 9025–9033.
[77] T. Kuznetsova, B. Kvamme, Grand canonical molecular dynamics for tip4p water systems, Mol. Phys. 97 (3) (1999) 423–431.
[78] H. Eslami, F. Müller-Plathe, Molecular dynamics simulation in the grand canonical ensemble, J. Comput. Chem. 28 (2007) 1763–1773.
[79] S. Chempath, L.A. Clark, R.Q. Snurr, Two general methods for grand canonical ensemble simulation of molecules with internal flexibility, J. Chem. Phys. 118 (2003) 7635–7643.
[80] M. Lsal, W.R. Smith, J. Kolafa, Molecular simulations of aqueous electrolyte solubility: 1. the expanded-ensemble osmotic molecular dynamics method for the solution phase, J. Phys. Chem. B 109 (2005) 12956–12965.
[81] R. Delgado-Buscalioni, J. Sablić, M. Praprotnik, Open boundary molecular dynamics, Eur. Phys. J. Spec. Top. 224 (2015) 2331–2349.
[82] J. Sablić, M. Praprotnik, R. Delgado-Buscalioni, Open boundary molecular dynamics of sheared star-polymer melts, Soft Matter 12 (2016) 2416–2439.
[83] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nature Struct. Biol. 9 (2002) 646–652.
[84] G.S. Ayton, W.G. Noid, G.A. Voth, Multiscale modeling of biomolecular systems: In serial and in parallel, Curr. Opin. Struct. Biol. 17 (2007) 192–198.
[85] M. Praprotnik, L. Delle Site, K. Kremer, Multiscale simulation of soft matter: From scale bridging to adaptive resolution, Annu. Rev. Phys. Chem. 59 (2008) 545–571.
[86] W. Tschöp, K. Kremer, O. Hahn, J. Batoulis, T. Bürger, Simulation of polymer melts. ii. from coarse-grained models back to atomistic description, Acta Polym. 49 (1998) 75–79.
[87] B. Hess, S. León, N. van der Vegt, K. Kremer, Long time atomistic polymer trajectories from coarse grained simulations: Bisphenol-a polycarbonate, Soft Matter 2 (2006) 409–414.
[88] T.A. Wassenaar, H.I. Ingólfsson, R.A. Böckmann, D.P. Tieleman, S.J. Marrink, Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations, J. Chem. Theory Comput. 11 (2015) 2144–2155.
[89] M. Praprotnik, L. Delle Site, K. Kremer, A macromolecule in a solvent: Adaptive resolution molecular dynamics simulation, J. Chem. Phys. 126 (2007) 134902.
[90] M. Praprotnik, L. Delle Site, K. Kremer, Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly, J. Chem. Phys. 123 (2005) 224106.
[91] M. Praprotnik, S. Poblete, K. Kremer, Statistical physics problems in adaptive resolution computer simulations of complex fluids, J. Stat. Phys. 145 (2011) 946–966.
[92] M. Praprotnik, L. Delle Site, Multiscale molecular modeling, in: L. Monticelli, E. Salonen (Eds.), Biomolecular Simulations: Methods and Protocols, in: Methods in Molecular Biology, vol. 924, Springer Science+Business Media, New York, 2013, pp. 567–583.
[93] T.P. Straatsma, J.A. McCammon, Computational alchemy, Annu. Rev. Phys. Chem. 43 (1992) 407–435.
[94] R.W. Zwanzig, High-temperature equation of state by a perturbation method. 1. nonpolar gases, J. Chem. Phys. 22 (1954) 1420.
[95] D.L. Beveridge, F.M. DiCapua, Free energy via molecular simulation: Applications to chemical and biomolecular systems, Annu. Rev. Biophys. Biophys. Chem. 18 (1989) 431–492.
[96] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with namd, J. Comput. Chem. 26 (2005) 1781–1802.
[97] M. Praprotnik, L. Delle Site, K. Kremer, Adaptive resolution scheme for efficient hybrid atomistic-mesoscale molecular dynamics simulations of dense liquids, Phys. Rev. E 73 (2006) 066701.
[98] M. Praprotnik, K. Kremer, L. Delle Site, Adaptive molecular resolution via a continuous change of the phase space dimensionality, Phys. Rev. E 75 (2007) 017701.
[99] M. Praprotnik, K. Kremer, L. Delle Site, Fractional dimensions of phase space variables: a tool for varying the degrees of freedom of a system in a multiscale treatment, J. Phys. A 40 (2007) F281–F288.
[100] L. Delle Site, Some fundamental problems for an energy-conserving adaptive-resolution molecular dynamics scheme, Phys. Rev. E 76 (2007) 047701.
[101] M. Praprotnik, S. Poblete, L. Delle Site, K. Kremer, Comment on "adaptive multiscale molecular dynamics of macromolecular fluids", Phys. Rev. Lett. 107 (2011) 099801.
[102] A.V. Ivlev, J. Bartnick, M. Heinen, C.-R. Du, V. Nosenko, H. Löwen, Statistical mechanics where newton's third law is broken, Phys. Rev. X 5 (2015) 011035.
[103] R. Potestio, S. Fritsch, P. Español, R. Delgado-Buscalioni, K. Kremer, R. Everaers, D. Donadio, Hamiltonian adaptive resolution simulation for molecular liquids, Phys. Rev. Lett. 110 (2013) 108301.
[104] R. Potestio, P. Español, R. Delgado-Buscalioni, R. Everaers, K. Kremer, D. Donadio, Monte Carlo adaptive resolution simulation of multicomponent molecular liquids, Phys. Rev. Lett. 111 (2013) 060601.
[105] C. Peter, K. Kremer, Multiscale simulation of soft matter systems - from the atomistic to the coarse-grained level and back, Soft Matter 5 (2009) 4357–4366.
[106] W.G. Noid, Perspective: Coarse-grained models for biomolecular systems, J. Chem. Phys. 139 (2013) 090901.
[107] R.L. Henderson, A uniqueness theorem for fluid pair correlation functions, Phys. Lett. 49A (1974) 197–198.

[108] A.K. Soper, Empirical Monte Carlo simulation of fluid structure, Chem. Phys. 202 (1996) 295–306.
[109] W. Tschöp, K. Kremer, J. Batoulis, T. Bürger, O. Hahn, Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates, Acta Polym. 49 (1998) 61–74.
[110] D. Reith, M. Pütz, F. Müller-Plathe, Deriving effective mesoscale potentials from atomistic simulations, J. Comput. Chem. 24 (2003) 1624–1636.
[111] A.P. Lyubartsev, A. Laaksonen, Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach, Phys. Rev. E 52 (1995) 3730–3737.
[112] A. Lyubartsev, A. Laaksonen, Determination of effective pair potentials from ab initio simulations: application to liquid water, Chem. Phys. Lett. 325 (2000) 15–21.
[113] S. Izvekov, M. Parrinello, C.B. Burnham, G.A. Voth, Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching, J. Chem. Phys. 120 (2004) 10896–10913.
[114] S. Izvekov, G.A. Voth, Multiscale coarse graining of liquid-state systems, J. Chem. Phys. 123 (2005) 134105.
[115] S. Izvekov, G.A. Voth, A multiscale coarse-graining method for biomolecular systems, J. Phys. Chem. B 109 (2005) 2469–2473.
[116] W.G. Noid, J.-W. Chu, G.S. Ayton, V. Krishna, S. Izvekov, G.A. Voth, A. Das, H.C. Andersen, The multiscale coarse-graining method. i. a rigorous bridge between atomistic and coarse-grained models, J. Chem. Phys. 128 (24) (2008) 244114.
[117] J.W. Mullinax, W.G. Noid, Generalized Yvon-Born-Green theory for molecular systems, Phys. Rev. Lett. 103 (2009) 198104.
[118] J.W. Mullinax, W.G. Noid, Extended ensemble approach for deriving transferable coarse-grained potentials, J. Chem. Phys. 131 (2009) 104110.
[119] M.S. Shell, The relative entropy is fundamental to thermodynamic ensemble optimization, J. Chem. Phys. 129 (2008) 144108.
[120] S.P. Carmichael, M.S. Shell, A new multiscale algorithm and its application to coarse-grained peptide models for self-assembly, J. Phys. Chem. B 116 (2012) 8383–8393.
[121] T. Foley, M.S. Shell, W.G. Noid, The impact of resolution upon entropy and information in coarse-grained models, J. Chem. Phys. 143 (2015) 243104.
[122] T.F. Nonnenmacher, Fractional integral and diferential equations for a class of Levi-type probability densities, J. Phys. A: Math. Gen. 23 (1990) L697S–L700S.
[123] R. Hilfer (Ed.), Applications of Fractional Calculus in Physics, World Scientific Publishing, Co. Pte. Ltd., Singapore, 2000.
[124] K. Cotrill-Shepherd, M. Naber, Fractional differential forms, J. Math. Phys. 42 (2001) 2203–2212.
[125] V.E. Tarasov, Fractional generalization of Liouville equations, Chaos 14 (2004) 123–127.
[126] V.E. Tarasov, Fractional systems and fractional Bogoliubov hiearchy equations, Phys. Rev. E 71 (2005) 011102.
[127] D. Janežič, M. Praprotnik, F. Merzel, Molecular dynamics integration and molecular vibrational theory: I. New Symplectic Integrators, J. Chem. Phys. 122 (2005) 174101.
[128] S. Matysiak, C. Clementi, M. Praprotnik, K. Kremer, L. Delle Site, Modeling diffusive dynamics in adaptive resolution simulation of liquid water, J. Chem. Phys. 128 (2008) 024503.
[129] S. Poblete, M. Praprotnik, K. Kremer, L. Delle Site, Coupling different levels of resolution in molecular simulations, J. Chem. Phys. 132 (2010) 114101.
[130] J. Zavadlav, R. Podgornik, M. Praprotnik, Adaptive resolution simulation of a DNA molecule in salt solution, J. Chem. Theory Comput. 11 (2015) 5035–5044.
[131] P. Español, R. Delgado-Buscalioni, R. Everaers, R. Potestio, D. Donadio, K. Kremer, Statistical mechanics of hamiltonian adaptive resolution simulations, J. Chem. Phys. 142 (2015) 064115.
[132] R. Delgado-Buscalioni, Thermodynamics of adaptive molecular resolution, Phil. Trans. R. Soc. A 374 (2016) 20160152.
[133] S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, K. Kremer, Adaptive resolution molecular dynamics simulation through coupling to an internal particle reservoir, Phys. Rev. Lett. 108 (2012) 170602.
[134] E.M. Kotsalis, J.H. Walther, P. Koumoutsakos, Control of density fluctuations in atomistic-continuum simulations of dense liquids, Phys. Rev. E 76 (2007) 016709.
[135] E.M. Kotsalis, J.H. Walther, E. Kaxiras, P. Koumoutsakos, Control algorithm for multiscale flow simulations of water, Phys. Rev. E 79 (2009) 045701.
[136] S. Bevc, C. Junghans, K. Kremer, M. Praprotnik, Adaptive resolution simulation of salt solutions, New J. Phys. 15 (2013) 105007.
[137] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6, J. Comput. Chem. 25 (2004) 1656–1676.
[138] Y. Duan, C. Wu, S. Chowdhury, M. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (2003) 1999–2012.
[139] M. Praprotnik, S. Matysiak, L. Delle Site, K. Kremer, C. Clementi, Adaptive resolution simulation of liquid water, J. Phys.: Condens. Matter 19 (2007) 292201.
[140] A. Nagarajan, C. Junghans, S. Matysiak, Multiscale simulation of liquid water using a four-to-one mapping for coarse-graining, J. Chem. Theory Comput. 9 (2013) 5168–5175.
[141] J. Zavadlav, M.N. Melo, S.J. Marrink, M. Praprotnik, Adaptive resolution simulation of an atomistic protein in martini water, J. Chem. Phys. 140 (2014) 054114.
[142] J. Zavadlav, M.N. Melo, A.V. Cunha, A.H. de Vries, S.J. Marrink, M. Praprotnik, Adaptive resolution simulation of martini solvents, J. Chem. Theory Comput. 10 (2014) 2591–2598.
[143] J. Zavadlav, M.N. Melo, S.J. Marrink, M. Praprotnik, Adaptive resolution simulation of polarizable supramolecular coarse-grained water models, J. Chem. Phys. 142 (2015) 244118.
[144] J. Zavadlav, R. Podgornik, M.N. Melo, S.J. Marrink, M. Praprotnik, Adaptive resolution simulation of an atomistic dna molecule in MARTINI salt solution, Eur. Phys. J. Spec. Top. 225 (2016) 1595–1607.
[145] J. Zavadlav, S.J. Marrink, M. Praprotnik, Adaptive resolution simulation of supramolecular water: The concurrent making, breaking, and remaking of water bundles, J. Chem. Theory Comput. 12 (2016) 4138–4145.
[146] S.J. Marrink, A.H. de Vries, A.E. Mark, Coarse grained model for semiquantitative lipid simulations, J. Phys. Chem. B 108 (2004) 750–760.
[147] S.J. Marrink, D.P. Tieleman, Perspective on the martini model, Chem. Soc. Rev. 42 (2013) 6801–6822.
[148] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The martini force field: Coarse grained model for biomolecular simulations, J. Phys. Chem. B 111 (2007) 7812–7824.
[149] K. Kreis, A. Fogarty, K. Kremer, R. Potestio, Advantages and challenges in coupling an ideal gas to atomistic models in adaptive resolution simulations, Eur. Phys. J. Spec. Top. 224 (2015) 2289–2304.
[150] U. Alekseeva, R.G. Winkler, G. Sutmann, Hydrodynamics in adaptive resolution particle simulations: Multiparticle collision dynamics, J. Comput. Phys. 314 (2016) 14–34.
[151] N.D. Petsev, L.G. Leal, M.S. Shell, Hybrid molecular-continuum simulations using smoothed dissipative particle dynamics, J. Chem. Phys. 142 (2015) 044101.
[152] A.C. Fogarty, R. Potestio, K. Kremer, Adaptive resolution simulation of a biomolecule and its hydration shell: Structural and dynamical properties, J. Chem. Phys. 142 (2015) 195101.

[153] F. Stanzione, A. Jayaraman, Hybrid atomistic and coarse-grained molecular dynamics simulations of polyethylene glycol (peg) in explicit water, J. Phys. Chem. B 120 (2016) 4160–4173.
[154] O.M. Szklarczyk, N.S. Bieler, P.H. Hünenberger, W.F. van Gunsteren, Flexible boundaries for multiresolution solvation: An algorithm for spatial multiscaling in molecular dynamics simulations, J. Chem. Theory Comput. 11 (2015) 5447–5463.
[155] K. Kreis, R. Potestio, K. Kremer, A.C. Fogarty, Adaptive resolution simulations with self-adjusting high-resolution regions, J. Chem. Theory Comput. 12 (2016) 4067–4081.
[156] C. Junghans, M. Praprotnik, K. Kremer, Transport properties controlled by a thermostat: An extended dissipative particle dynamics thermostat, Soft Matter 4 (2008) 156–161.
[157] R. Zwanzig, Nonequilibrium Statistical Mechanics, Oxford University Press, 2001.
[158] H. Mori, Transport, collective motion, and brownian motion, Progr. Theoret. Phys. 33 (3) (1965) 423–455.
[159] S. Izvekov, G.A. Voth, Modeling real dynamics in the coarse-grained representation of condensed phase systems, J. Chem. Phys. 125 (2006) 151101.
[160] C. Hijon, P. Espanol, E. Vanden-Eijnden, R. Delgado-Buscalioni, Mori-Zwanzig formalism as a practical computational tool, Faraday Discuss. 144 (2010) 301–322.
[161] Z. Li, X. Bian, B. Caswell, G.E. Karniadakis, Construction of dissipative particle dynamics models for complex fluids via the mori-zwanzig formulation, Soft Matter 10 (2014) 8659–8672.
[162] Z. Li, X. Bian, X. Li, G.E. Karniadakis, Incorporation of memory effects in coarse-grained modeling via the mori-zwanzig formalism, J. Chem. Phys. 143 (24) (2015) 243128.
[163] Z. Li, X. Bian, X. Yang, G.E. Karniadakis, A comparative study of coarse-graining methods for polymeric fluids: Mori-Zwanzig vs. iterative Boltzmann inversion vs. stochastic parametric optimization, J. Chem. Phys. 145 (2016) 044102.
[164] T. Soddemann, B. Dünweg, K. Kremer, Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations, Phys. Rev. E 68 (2003) 046702.
[165] P.J. Hoogerbrugge, J.M.V.A. Koelman, Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics, Europhys. Lett. 19 (1992) 155–160.
[166] P. Español, P. Warren, Statistical mechanics of dissipative particle dynamics, Europhys. Lett. 30 (1995) 191–196.
[167] J.D. Halverson, T. Brandes, O. Lenz, A. Arnold, S. Bevc, V. Starchenko, K. Kremer, T. Stuehn, D. Reith, Espresso++: A modern multiscale simulation package for soft matter systems, Comput. Phys. Comm. 184 (2013) 1129–1149.
[168] S. Fritsch, C. Junghans, K. Kremer, Structure formation of toluene around c60: Implementation of the adaptive resolution scheme (adress) into gromacs, J. Chem. Theory Comput. 8 (2012) 398–403.
[169] C.F. Abrams, Concurrent dual-resolution Monte Carlo simulation of liquid methane, J. Chem. Phys. 123 (2005) 234101.
[170] A. Heyden, H. Lin, D.G. Truhlar, Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations, J. Phys. Chem. B 111 (2007) 2231–2241.
[171] B. Ensing, S.O. Nielsen, P.B. Moore, M.L. Klein, M. Parrinello, Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics, J. Chem. Theory Comput. 3 (2007) 1100.
[172] S.O. Nielsen, P.B. Moore, B. Ensing, Adaptive multiscale molecular synamics of macromolecular fluids, Phys. Rev. Lett. 105 (2010) 237802.
[173] K. Kreis, R. Potestio, The relative entropy is fundamental to adaptive resolution simulations, J. Chem. Phys. 145 (4) (2016) 044104.
[174] M. Heidari, R. Cortes-Huerto, D. Donadio, R. Potestio, Accurate and general treatment of electrostatic interaction in hamiltonian adaptive resolution simulations, Eur. Phys. J. Spec. Top. 225 (2016) 1505–1526.
[175] R. Everaers, Thermodynamic translational invariance in concurrent multiscale simulations of liquids, Eur. Phys. J. Spec. Top. 225 (2016) 1483–1503.
[176] M. Christen, W.F. van Gunsteren, Multigraining: An algorithm for simultaneous fine-grained and coarse-grained simulation of molecular systems, J. Chem. Phys. 124 (2006) 154106.
[177] L. Delle Site, C.F. Abrams, A. Alavi, K. Kremer, Polymers near metal surfaces: Selective adsorption and global conformations, Phys. Rev. Lett. 89 (2002) 156103.
[178] M. Neri, C. Anselmi, M. Cascella, A. Maritan, P. Carloni, Coarse-grained model of proteins incorporating atomistic detail of the active site, Phys. Rev. Lett. 95 (2005) 218102.
[179] A. Heyden, D.G. Truhlar, Conservative algorithm for an adaptive change of resolution in mixed atomistic/coarse-grained multiscale simulations, J. Chem. Theory Comput. 4 (2008) 217.
[180] J.H. Park, A. Heyden, Solving the equations of motion for mixed atomistic and coarse-grained systems, Mol. Simul. 35 (10–11) (2009) 962–973.
[181] M. Böckmann, N. Doltsinis, D. Marx, Adaptive switching of interaction potentials in the time domain: An extended lagrangian approach tailored to transmute force field to qm/MM Simulations and Back, J. Chem. Theory Comput. 11 (2015) 2429–2439.
[182] A.J. Rzepiela, M. Louhivuori, C. Peter, S.J. Marrink, Hybrid simulations: Combining atomistic and coarse-grained force fields using virtual sites, Phys. Chem. Chem. Phys. 13 (2011) 10437–10448.
[183] T.A. Wassenaar, H.I. Ingólfsson, M. Priess, S.J. Marrink, L.V. Schaefer, Mixing MARTINI: Electrostatic coupling in hybrid atomistic - coarse-grained biomolecular simulations, J. Phys. Chem. B 117 (2013) 3516–3530.
[184] P. Sokkar, S.M. Choi, Y.M. Rhee, Simple method for simulating the mixture of atomistic and coarse-grained molecular systems, J. Chem. Theory Comput. 9 (2013) 3728–3739.
[185] P. Sokkar, E. Boulanger, W. Thiel, E. Sanchez-Garcia, Hybrid quantum mechanics/molecular mechanics/coarse grained modeling: A triple-resolution approach for biomolecular systems, J. Chem. Theory Comput. 11 (2015) 1809–1818.
[186] M. Orsi, W. Ding, M. Palaiokostas, Direct mixing of atomistic solutes and coarse-grained water, J. Chem. Theory Comput. 10 (2014) 4684–4693.
[187] M.R. Machado, P.D. Dans, S. Pantano, A hybrid all-atom/coarse grain model for multiscale simulations of DNA, Phys. Chem. Chem. Phys. 13 (2011) 18134–18144.
[188] L. Darr, A. Tek, M. Baaden, S. Pantano, Mixing atomistic and coarse grain solvation models for md simulations: Let WT4 Handle the Bulk, J. Chem. Theory Comput. 8 (10) (2012) 3880–3894.
[189] H.C. Gonzalez, L. Darré, S. Pantano, Transferable mixing of atomistic and coarse-grained water models, J. Phys. Chem. B 117 (2013) 14438–14448.
[190] M.R. Machado, S. Pantano, Exploring LacIDNA dynamics by multiscale simulations using the SIRAH force field, J. Chem. Theory Comput. 11 (2015) 5012–5023.
[191] A. Chaimovich, C. Peter, K. Kremer, Relative resolution: A hybrid formalism for fluid mixtures, J. Chem. Phys. 143 (24) (2015) 243107.
[192] S. Riniker, W.F. van Gunsteren, Mixing coarse-grained and fine-grained water in molecular dynamics simulations of a single system, J. Chem. Phys. 137 (2012) 044120.
[193] S. Artemova, S. Redon, Adaptively restrained particle simulations, Phys. Rev. Lett. 109 (2012) 190201.
[194] K. Kreis, M.E. Tuckerman, D. Donadio, K. Kremer, R. Potestio, From classical to quantum and back: A hamiltonian scheme for adaptive multiresolution classical/path-integral simulations, J. Chem. Theory Comput. 12 (7) (2016) 3030–3039.
[195] J.M. Boereboom, R. Potestio, D. Donadio, R.E. Bulo, Toward Hamiltonian adaptive QM/MM: Accurate solvent structures using many-body potentials, J. Chem. Theory Comput. 12 (2016) 3441.

[196] H. Wang, C. Schütte, L. Delle Site, Adaptive resolution simulation (adress): A smooth thermodynamic and structural transition from atomistic to coarse grained resolution and vice versa in a grand canonical fashion, J. Chem. Theoret. Comput. 8 (2012) 2878.
[197] H. Wang, C. Hartmann, C. Schütte, L. Delle Site, Grand-canonical-like molecular-dynamics simulations by using an adaptive-resolution technique, Phys. Rev. X 3 (2013) 011018.
[198] A. Agarwal, H. Wang, C. Schütte, L. Delle Site, Chemical potential of liquids and mixtures via adaptive resolution simulation, J. Chem. Phys. 141 (2014) 034102.
[199] A. Agarwal, J. Zhu, C. Hartmann, H. Wang, L. Delle Site, Molecular dynamics in a grand ensemble: Bergmann-Lebowitz model and adaptive resolution simulation, New. J. Phys. 17 (2015) 083042.
[200] E.G. Flekkoy, R. Delgado-Buscalioni, P.V. Coveney, Flux boundary conditions in particle simulations, Phys. Rev. E 72 (2005) 026703.
[201] R. Delgado-Buscalioni, Tools for multiscale simulation of liquids using open molecular dynamics, Numer. Anal. Multiscale Comput. (2012) 145–166.
[202] C. Perego, F. Giberti, M. Parrinello, Chemical potential calculations in dense liquids using metadynamics, Eur. Phys. J. Spec. Top. 225 (2016) 1621–1628.
[203] R. Delgado-Buscalioni, P.V. Coveney, USHER: an algorithm for particle insertion in dense fluids, J. Chem. Phys. 119 (2003) 978–987.
[204] G. De Fabritiis, R. Delgado-Buscalioni, P.V. Coveney, Energy controlled insertion of polar molecules in dense fluids, J. Chem. Phys. 121 (2004) 12139.
[205] M.K. Borg, D.A. Lockerby, J.M. Reese, The fade mass-stat: A technique for inserting or deleting particles in molecular dynamics simulations, J. Chem. Phys. 140 (7) (2014) 074110.
[206] E.G. Flekkoy, G. Wagner, J. Feder, Hybrid model for combined particle and continuum dynamics, Europhys. Lett. 52 (2000) 271–276.
[207] R. Delgado-Buscalioni, A. Dejoan, Nonreflecting boundaries for ultrasound in fluctuating hydrodynamics of open systems, Phys. Rev. E 78 (2008) 046708.
[208] K.M. Mohamed, A.A. Mohamad, A review of the development of hybrid atomistic-scontinuum methods for dense fluids, Microfluid Nanofluid 8 (2010) 283–302.
[209] D. Mukherji, K. Kremer, Coil-globule-coil transition of pnipam in aqueous methanol: Coupling all-atom simulations to semi-grand canonical coarse-grained reservoir, Macromolecules 46 (22) (2013) 9158–9163.
[210] R. Klein, Comments on "open boundary molecular dynamics" by R. Delgado-Buscalioni, J. Sablić, and M. Praprotnik, Eur. Phys. J. Spec. Top. 224 (2015) 2509–2510.
[211] R. Delgado-Buscalioni, J. Sablić, M. Praprotnik, Reply to comments by R. Klein on "Open boundary molecular dynamics", Eur. Phys. J. Spec. Top. 224 (2015) 2511–2513.
[212] S.T. O'Connell, P.A. Thompson, Molecular dynamics-continuum hybrid computations: A tool for studying complex fluid flows, Phys. Rev. E 52 (1995) R5792.
[213] N.G. Hadjiconstantinou, Combining atomistic and continuum simulations of contact-line motion, Phys. Rev. E 59 (1999) 2475–2478.
[214] X. Nie, S. Chen, M.O. Robbins, Hybrid continuum-atomistic simulation of singular corner flow, Phys. Fluids 16 (10) (2004) 3579–3591.
[215] T. Werder, J.H. Walther, P. Koumoutsakos, Hybrid atomistic-continuum method for the simulation of dense fluid flows, J. Comput. Phys. 205 (2005) 373–390.
[216] D.A. Fedosov, G.E. Karniadakis, Triple-decker: Interfacing atomistic-mesoscopic-continuum flow regimes, J. Comput. Phys. 228 (2009) 1157–1171.
[217] J.H. Walther, M. Praprotnik, E.M. Kotsalis, P. Koumoutsakos, Multiscale simulation of water flow past a {C540} fullerene, J. Comput. Phys. 231 (7) (2012) 2677–2681.
[218] G.D. Fabritiis, R. Delgado-Buscalioni, P.V. Coveney, Multiscale modeling of liquids with molecular specificity, Phys. Rev. Lett. 97 (2006) 134501.
[219] R. Delgado-Buscalioni, K. Kremer, M. Praprotnik, Concurrent triple-scale simulation of molecular liquids, J. Chem. Phys. 128 (2008) 114110.
[220] W. E, B. Enquist, X.T. Li, W.Q. Ren, E. Vanden-Eijden, Heterogeneous multiscale methods: a review, CiCP 2 (2007) 367–450.
[221] R. Delgado-Buscalioni, P.V. Coveney, Continuum-particle hybrid coupling for mass, momentum, and energy transfers in unsteady fluid flow, Phys. Rev. E 67 (2003) 046704.
[222] R. Delgado-Buscalioni, P.V. Coveney, Hybrid molecular-continuum fluid dynamics, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 362 (2004) 1639–1654.
[223] R. Delgado-Buscalioni, P.V. Coveney, G.D. Riley, R.W. Ford, Hybrid molecular-continuum fluid models: implementation within a general coupling framework, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 363 (2005) 1975–1985.
[224] R. Delgado-Buscalioni, G. De Fabritiis, Embedding molecular dynamics within fluctuating hydrodynamics in multiscale simulations of liquids, Phys. Rev. E 76 (2007) 036709.
[225] R. Delgado-Buscalioni, K. Kremer, M. Praprotnik, Coupling atomistic and continuum hydrodynamics through a mesoscopic model: application to liquid water, J. Chem. Phys. 131 (2009) 244107.
[226] P. Koumoutsakos, Multiscale flow simulations using particles, Annu. Rev. Fluid Mech. 37 (2005) 457.
[227] P. Español, M. Revenga, Smoothed dissipative particle dynamics, Phys. Rev. E 67 (2003) 026705.
[228] L.B. Lucy, A numerical approach to the testing of the fission hypothesis, Astron. J. 82 (1977) 1013–1024.
[229] R.A. Gingold, J.J. Monaghan, Smoothed particle hydrodynamics - theory and application to non-spherical stars, Mon. Not. R. Astron. Sci. 181 (1977) 375–389.
[230] A. Malevanets, R. Kapral, Mesoscopic model for solvent dynamics, J. Chem. Phys. 110 (1999) 8605.
[231] A. Scukins, D. Nerukh, E. Pavlov, S. Karabasov, A. Markesteijn, Multiscale molecular dynamics/hydrodynamics implementation of two dimensional "mercedes benz" water model, Eur. Phys. J. Spec. Top. 224 (12) (2015) 2217–2238.
[232] R. Feynman, A. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, 1965.
[233] D. Scharf, G. Martyna, M. Klein, Structure and energetics of fluid para-hydrogen, Low Temp. Phys. 19 (1993) 364.
[234] A. Wallqvist, B. Berne, Path integral simulation of pure water, Chem. Phys. Lett. 117 (1985) 214.
[235] M. Tuckerman, Path integration via molecular dynamics, in: J. Grotendorst, D. Marx, A. Muramatsu (Eds.), Qunatum Simulations of Complex Many-Body Systems: From Theory to Algorithms, in: NIC Series, vol. 10, 2002, 2014, pp. 169–189.
[236] Q. Wang, K. Johnson, J. Broughton, Path integral grand canonical Monte Carlo, J. Chem. Phys. 107 (1997) 5108.
[237] G. Garberoglio, Boltzmann bias grand canonical Monte Carlo, J. Chem. Phys. 128 (2008) 134109.
[238] A. Poma, L. Delle Site, Classical to path-integral adaptive resolution in molecular simulation: Towards a smooth quantum-classical coupling, Phys. Rev. Lett. 104 (2010) 250201.
[239] A. Poma, L. Delle Site, Adaptive resolution simulation of liquid para-hydrogen: Testing the robustness of the quantum-classical adaptive coupling, Phys. Chem. Chem. Phys. 13 (2011) 10510.
[240] R. Potestio, L. Delle Site, Quantum locality and equilibrium properties in low-temperature parahydrogen: A multiscale simulation study, J. Chem. Phys. 136 (2012) 054101.
[241] A. Agarwal, L. Delle Site, Path integral molecular dynamics within the grand canonical-like adaptive resolution technique: Simulation of liquid water, J. Chem. Phys. 143 (2015) 094102.
[242] A. Agarwal, L. Delle Site, Grand-canonical adaptive resolution centroid molecular dynamics: Implementation and application, Comput. Phys. Comm. 206 (2016) 26.
[243] R. Potestio, Computer simulation of particles with position-dependent mass, Eur. Phys. J. B 87 (2014) 245.

[244] A. Warshel, M. Levitt, Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, J. Mol. Biol. 103 (1976) 227.
[245] M. Field, P. Bash, M. Karplus, A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations, J. Comput. Chem. 11 (1990) 700.
[246] H. Senn, W. Thiel, QM/MM Methods for Biomolecular Systems, Angew. Chem. Int. Ed. 48 (2009) 1198.
[247] E. Brunk, U. Rothlisberger, Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states, Chem. Rev. 115 (2015) 6217.
[248] H. Watanabe, T. Kubar, M. Elstner, Size-Consistent multipartitioning QM/MM: A stable and efficient adaptive QM/MM, J. Chem. Theory Comput. 10 (2014) 4242.
[249] R. Bulo, B. Ensing, J. Sikkema, L. Visscher, Toward a practical method for adaptive QM/MM Simulations, J. Chem. Theory. Comput. 5 (2009) 2212.
[250] S. Nielsen, R. Bulo, P. Moore, B. Ensing, Recent progress in adaptive multiscale molecular dynamics simulation of soft matter, Phys. Chem. Chem. Phys. 12 (2010) 12401.
[251] S. Pezeshki, H. Lin, Adaptive-partitioning QM/MM dynamics simulations: 4. Solvent molecules entering and leaving protein binding sites, J. Chem. Theory Comput. 10 (2014) 4765.
[252] S. Pezeshki, C. Davis, A. Heyden, H. Lin, Adaptive-partitioning QM/MM dynamics simulations: 4. Proton hopping in bulk water, J.Chem. Theory Comput. 10 (2014) 4765.
[253] N. Bernstein, C. Varnai, I. Solt, S. Winfield, M. payne, I. Simon, M. Fuxreiter, G. Csany, QM/MM simulation of liquid water with an adaptive quantum region, Phys. Chem. Chem. Phys. 14 (2012) 646.
[254] L. Mones, A. Jones, A. Götz, T. Laino, R. Walker, B. Leimkuhler, G. Csany, N. Bernstein, The adaptive buffered force QM/MM Method in the CP2K and Amber Software Packages, J. Comput. Chem. 36 (2015) 633.
[255] E. E., R. Dreizler, Density Functional Theory, Springer, 2011.
[256] W. Parr, R.G. Yang, Density-Functional Theory of Atoms and Molecules, Oxford University Press, 1989.
[257] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. Rev. 136 (1964) B864.
[258] W. Kohn, L. Sham, Self-Consistent equations including exchange and correlation effects, Phys. Rev. 140 (1965) A1133.
[259] D. Mermin, Thermal properties of the inhomogeneous electron gas, Phys. Rev. 137 (1965) A1441.
[260] J. Capitani, R. Nalewajski, R. Parr, Non-born-Oppenheimer density functional theory of molecular systems, J. Chem. Phys. 76 (1981) 568.
[261] J. Perdew, R. Parr, M. Levy, J. Balduz, Density-functional theory for fractional number: derivative discontinuites of the energy, Phys. Rev. Lett. 49 (1982) 1691.
[262] W. Kohn, A. Becke, R. Parr, Density functional theory of electronic structure, J. Phys. Chem. 100 (1996) 12974.
[263] W. Schneider, A. Auer, Constant chemical potential approach for quantum chemical calculations in electrocatalysis, Beilstein J. Nanotechnol. 5 (2014) 668.
[264] J. Kirkwood, Statistical mechanics of fluid mixtures, J. Phys. Chem. 3 (1935) 300.
[265] S. Jacobi, R. Baer, Variational grand canonical electronic structure method for open systems, J. Chem. Phys. 123 (2005) 044112.
[266] R. Sedgewick, D. Scalapino, R. Sugar, L. Capriotti, Grand canonical and canonical ensemble expectation values from quantum Monte Carlo simulations, Phys. Rev. B 68 (2003) 045120.
[267] S. Jacobi, R. Baer, Variational grand-canonical electronic structure of $Li+Li$ at $\approx 104K$ with second-order perturbation theory corrections, Th. Chem. Acc. 131 (2012) 1113.
[268] A. Popadić, J.H. Walther, P. Koumoutsakos, M. Praprotnik, Continuum simulations of water flow in carbon nanotube membranes, New J. Phys. 16 (8) (2014) 082001.
[269] A. Popadić, M. Praprotnik, P. Koumoutsakos, J.H. Walther, Continuum simulations of water flow past fullerene molecules, Eur. Phys. J. Spec. Top. 224 (2015) 2321–2330.
[270] A.W. Lees, S.F. Edwards, The computer study of transport processes under extreme conditions, J. Phys. C Solid State 5 (15) (1972) 1921.
[271] H. Wang, C. Schütte, G. Ciccotti, L. Delle Site, Exploring the conformational dynamics of alanine dipeptide in solution subjected to an external electric field: a nonequilibrium molecular dynamics simulation, J. Chem. Theory Comput. 10 (2014) 1376–1386.
[272] G. Ciccotti, G. Jacucci, Direct computation of dynamical response by molecular dynamics: The mobility of a charged Lennard-Jones particle, Phys. Rev. Lett. 35 (1975) 789.
[273] G. Ciccotti, G. Jacucci, I. McDonald, "Thought-experiments" by molecular dynamics, J. Stat. Phys. 21 (1979) 1.
[274] S. Orlandini, S. Meloni, G. Ciccotti, Hydrodynamics from statistical mechanics: combined dynamical-nemd and conditional sampling to relax an interface between two immiscible liquids, Phys. Chem. Chem. Phys. 13 (2011) 13177.
[275] M. Ferrario, S. Bonella, G. Ciccotti, On the establishment of thermal diffusion in binary Lennard-Jones liquids, Eur. Phys. J. Spec. Top. 225 (2016) 1629–1642.
[276] K. Kreis, R. Potestio, K. Kremer, A. Fogarty, Adaptive resolution simulations with self-adjusting high-resolution regions, J. Chem. Theory Comput. 12 (2016) 4067.
[277] C. Van den Broeck, S. Sasa, U. Seifert, Focus on stochastic thermodynamics, New. J. Phys. 18 (2016) 020401.
[278] A. Isihara, The gibbs-Bogoliubov inequality, J. Phys. A 1 (1968) 539.
[279] B.J. Alder, T.E. Wainwright, Introduction to molecular dynamics simulation, J. Chem. Phys. 31 (1959) 459.
[280] H.F. Trotter, On the product of semi-groups of operators, Proc. Amer. Math. Soc. 10 (1959) 545–551.
[281] G. Strang, On the construction and comparison of difference schemes, SIAM J. Numer. Anal. 5 (1968) 506–517.
[282] M. Praprotnik, D. Janežič, Molecular dynamics integration meets standard theory of molecular vibrations, J. Chem. Inf. Model 45 (2005) 1571–1579.
[283] M. Praprotnik, D. Janežič, Molecular dynamics integration and molecular vibrational theory: II. Simulation of Non-Linear Molecules, J. Chem. Phys. 122 (2005) 174102.
[284] M. Praprotnik, D. Janežič, Molecular dynamics integration and molecular vibrational theory: III. The Infrared Spectrum of Water, J. Chem. Phys. 122 (2005) 174103.
[285] B. Leimkuhler, C. Matthews, Robust and efficient configurational molecular sampling via Langevin dynamics, J. Chem. Phys. 138 (2013) 174102.
[286] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon Press, 1989.
[287] D.P. Landau, K. Binder, A Guide to Monte Carlo Simulations in Statistical Physics, Cambridge University Press, 2013.
[288] K. Binder, Applications of Monte Carlo methods to statistical physics, Rep. Progr. Phys. 60 (1997) 487–559.
[289] H. Kleinert, Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets, World Scientific, 2006.
[290] I. Craig, D. Manolopoulos, Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics, J. Chem. Phys. 121 (2004) 3368.
[291] S. Habershon, D. Manolopoulos, T. Markland, T. Miller III, Ring polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space, Annu. Rev. Phys. Chem. 64 (2013) 387.
[292] J. Cao, G. Voth, A new perspective on quantum time correlation functions, J. Chem. Phys. 99 (1993) 10070.

[293] T. Hone, P. Rossky, G. Voth, A comparative study of imaginary time path integral based methods for quantum dynamics, J. Chem. Phys. 124 (2006) 154103.
[294] M. Ceriotti, M. Parrinello, T. Markland, D. Manolopoulos, Efficient stochastic thermostatting of path integral molecular dynamics, J. Chem. Phys. 133 (2010) 124104.
[295] R. Kubo, Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems, J. Phys. Soc. Japan 12 (1957) 570.
[296] R. Kubo, M. Toda, N. Hashitsume, Statistical Physics II: Nonequilibrium Statistical Mechanics, Springer, New York, 1985.
[297] A. Witt, S. Ivanov, M. Shiga, H. Forbert, D. Marx, On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy, J. Chem. Phys. 130 (2009) 194510.
[298] M. Pavese, S. Jang, G. Voth, Centroid molecular dynamics: a quantum dynamics method suitable for the parallel computer, Parallel Comput. 26 (2000) 1025.
[299] M. Perez, M. Tuckerman, M. Muser, A comparative study of the centroid and ring-polymer molecular dynamics methods for approximating quantum time correlation functions from path integrals, J. Chem. Phys. 130 (2009) 184105.
[300] S. A., N. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, Dover, 1996.
[301] W. Kohn, Density functional theory and density matrix method scaling linearly with the number of atoms, Phys. Rev. Lett. 76 (1996) 3168.
[302] R. Friesner, B. Guallar, Ab initio quantum mrchanical and mixed quantum mechanical/molecular mechanics (QM/MM) Methods for Studying Enzymatic Catalysis, Annu. Rev. Phys. Chem. 56 (2005) 389.
[303] A. Warshel, Computer simulation of enzyme catalysis: methods, progress, and insights, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 425.
[304] Y. Zhao, D. Truhlar, A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions, J. Chem. Phys. 125 (2006) 194101.
[305] B. Janesko, Rung 3.5 density functionals: another step on Jacob's ladder, Int. J. Quantum Chem. 113 (2013) 83.
[306] T. Gould, Beyond the random phase approximation on the cheap: Improved correlation energies with the efficient "radial exchange hole" kernel, J. Chem. Phys. 137 (2012) 111101.
[307] J. Perdew, S. Kurth, Density functionals for non-relativistic coulomb systems in the new century, in: C. Fiolhais, F. Noqueira, M.A.L. Marques (Eds.), A Primer in Density Functional Theory, in: Lecture Notes in Physics, vol. 620, Springer-Verlag, Berlin Heidelberg, 2003, 1.
[308] A. Alavi, J. Kohanoff, M. Parrinello, D. Frenkel, *Ab initio* molecular dynamics with excited electrons, Phys. Rev. Lett. 73 (1994) 2599.
[309] R. Nesbet, Fractional occupation numbers in density-functional theory, Phys. Rev. A 56 (1997) 2665.
[310] W. Yang, Y. Zhang, P. Ayers, Degenerate ground states and a fractional number of electrons in density and reduced density matrix functional theory, Phys. Rev. Lett. 84 (2000) 5172.
[311] D. Alcoba, R. Bochicchio, G. Massacessi, L. Lain, A. Torre, Grand-canonical-ensemble representability problem for the one-electron reduced density matrix, Phys. Rev. A 75 (2007) 012509.
[312] J. Janak, Proof that $\frac{\partial e}{\partial n_i} = \epsilon$ in density-functional theory, Phys. Rev. B 18 (1978) 7165.
[313] A. Fetter, J. Waleka, Quantum Theory of Many-Particle Systems, Dover, 2003.
[314] J. Negele, H. Orland, Quantum Many-Particle Systems, Westviewpress, 1998.