Path integral molecular dynamics within the grand canonical-like adaptive resolution technique: Simulation of liquid water

Animesh Agarwal and Luigi Delle Site

Citation: The Journal of Chemical Physics **143**, 094102 (2015); doi: 10.1063/1.4929738
View online: http://dx.doi.org/10.1063/1.4929738
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/143/9?ver=pdfcov
Published by the AIP Publishing

Path integral molecular dynamics within the grand canonical-like adaptive resolution technique: Simulation of liquid water

Animesh Agarwal[a] and Luigi Delle Site[b]
Institute for Mathematics, Freie Universität Berlin, Berlin, Germany

(Received 18 June 2015; accepted 17 August 2015; published online 1 September 2015)

Quantum effects due to the spatial delocalization of light atoms are treated in molecular simulation via the path integral technique. Among several methods, Path Integral (PI) Molecular Dynamics (MD) is nowadays a powerful tool to investigate properties induced by spatial delocalization of atoms; however, computationally this technique is very demanding. The above mentioned limitation implies the restriction of PIMD applications to relatively small systems and short time scales. One of the possible solutions to overcome size and time limitation is to introduce PIMD algorithms into the Adaptive Resolution Simulation Scheme (AdResS). AdResS requires a relatively small region treated at path integral level and embeds it into a large molecular reservoir consisting of generic spherical coarse grained molecules. It was previously shown that the realization of the idea above, at a simple level, produced reasonable results for toy systems or simple/test systems like liquid parahydrogen. Encouraged by previous results, in this paper, we show the simulation of liquid water at room conditions where AdResS, in its latest and more accurate Grand-Canonical-like version (GC-AdResS), is merged with two of the most relevant PIMD techniques available in the literature. The comparison of our results with those reported in the literature and/or with those obtained from full PIMD simulations shows a highly satisfactory agreement. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4929738]

I. INTRODUCTION

The structure and dynamics of liquids consisting of molecules that contain light atoms (e.g., hydrogen) can be influenced by the quantum effects due to the delocalization of atoms in space. In simulation, such systems are treated by modeling the atoms of the molecules via the path integral formalism of Feynman.[1–3] In particular, liquid water is a typical subject of interest given its role in many fields.[4] As explained in more detail in Sec. III, the computational effort is massive because the number of interatomic interactions becomes much larger compared to the classical case. As a consequence, the size of the system and the simulation time affordable with standard computer resources are rather limited. For liquid water at room condition, a system of 500 molecules for a simulation time of 1-2 ns is usually considered already expensive. The limited size and simulation time may imply that particle number density fluctuations are arbitrarily suppressed and some systems cannot be treated if not at high computational price (e.g., solvation of a large molecule in water). An optimal complementary technique would consist of a Grand Canonical (GC)-like scheme where (local) properties can be calculated by employing a computationally affordable path integral (PI) simulation of a small open region which, in statistical and thermodynamic equilibrium, exchanges particles and energy with a reservoir acting at small computational cost. One possible implementation of a grand canonical-like Molecular Dynamics (MD) technique is the Adaptive Resolution Simulation scheme (AdResS)[5–7] in its most accurate version of GC-AdResS.[8–12] For the simplest version of AdResS, it was shown that for a toy system (liquid of tetrahedral molecules) the embedding of a PIMD technique into the scheme produced rather encouraging results;[13] such results were confirmed and empowered by the application to simple/test systems like liquid parahydrogen at low temperature.[14,15] Meanwhile, the increased accuracy and more solid conceptual framework of the adaptive scheme (GC-AdResS) allow for the study of more complex systems and the calculation of a larger number of properties than before.[9–12] In this perspective, this paper reports the technical implementation of two different approaches to PIMD, Refs. 16–18 and Refs. 19 and 20, into our GC-AdResS. We show its application to liquid water and report results about static and dynamic properties. The comparison with reference data is highly satisfactory and suggests that GC-AdResS, as a complementary method, may play an important role in future applications of PIMD (today not feasible with full PIMD simulations). One can think, for example, of solvation of a large molecules (e.g., fullerene in water) and look at possible quantum effects in the structure of the solvation shell. However, it must also be mentioned that in light of recent advances in the methodology of path integral calculations, the gain in efficiency with the (current) PI-AdResS scheme presented here needs to be compared with the gain of more advanced PIMD schemes. For example, two schemes, in particular, reduce the cost of path integral calculations by allowing a reduction in the number of beads. These are the path integral plus the generalized Langevin equation thermostat (PIGLET) of Manolopoulos and coworkers[19,21] and algorithms based on the Takahashi-Imada factorization (see, e.g., Refs. 22 and 23). However, such

[a]animesh@zedat.fu-berlin.de
[b]dellesite@fu-berlin.de

approaches do not work for the Ring Polymer MD (RPMD) scheme, employed later in this work for calculating equilibrium time correlation functions; in such case, the ring polymer contraction scheme would be appropriate.[20,24,25] This method is applicable to RPMD simulations and it leads to essentially classical (1 bead) numerical effort in the limit of large system size; therefore, it can be expected to be highly competitive with the current version of the method we propose. It must also be clear that the approaches reported above are integration techniques which are more efficient than the PIMD techniques that we have merged with AdResS in this paper; however, they may be merged with AdResS as well and thus take advantage of the grand canonical-like approach in reducing the number of degrees of freedom (even further). Moreover, a point that is certainly important is that GC-AdResS may be employed as a tool of analysis and study how the quantum effects change as a function of the size of the region treated at PI level. This would represent a novel type of analysis because it unequivocally defines the essential molecular degrees of freedom required for a given property[26] and thus, it allows to quantify how localized (possible) quantum effects (for the properties considered) are. The paper is organized as follows: Sec. II is dedicated to a summary of the relevant technical and conceptual characteristics of GC-AdResS, followed by Sec. III, which is dedicated to the description of the basic characteristics of the two PIMD methods employed in this study. The implementation of PIMD in GC-AdResS, for each of the two specific techniques used, is reported in Sec. III B. Sec. IV is divided into the subsections of (i) static and (ii) dynamic properties. In Subsection IV A, we report particle number density profiles, probability distributions, and radial distribution functions of the GC-AdResS simulation compared with results from full PIMD simulations. In Subsection IV B, we report the calculation of equilibrium time correlation functions compared, also in this case, with data obtained from full PIMD simulations. Finally, the conclusion is presented in Sec. V. The Appendix instead reports all technical data of the simulations so that the results can be reproduced/checked by other groups.

II. GC-ADRESS

In the original AdResS, the coupling idea is rather simple, that is, in a region of interest (the atomistic or high resolution region), all the molecular degrees of freedom are treated via molecular dynamics, while in a (larger) region of minor interest, only coarse-grained degrees of freedom are treated. The passage of a molecule from one region to another should be performed smoothly with hybrid dynamics in such a way that the atomistic and coarse-grained regions are not perturbed in a significant way. In order to do so, the space is divided into three regions: the atomistic (high resolution) region, the coarse-grained region, and an interfacial region where the atomistic degrees of freedom are transformed in coarse-grained and vice versa, we call this region hybrid region or transition region (see Fig. 1 and Ref. 27). The coupling is made via a space dependent force interpolation,

$$F_{\alpha\beta} = w(X_\alpha)w(X_\alpha)F^{atom}_{\alpha\beta} + [1 - w(X_\alpha)w(X_\alpha)]F^{cm}_{\alpha\beta}, \quad (1)$$

where α and β indicate two molecules, and $w(X_\alpha)$ and $w(X_\beta)$ indicate the interpolating (weighting) functions depending on the coordinate of the center of mass of the molecules X_α and X_β,

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

where $d_{AT}$and d_Δ are the size of atomistic and hybrid regions, respectively. $F^{atom}_{\alpha\beta}$ is the force in the atomistic region, which is derived from atomistic interactions, and $F^{cm}_{\alpha\beta}$ is the force in the coarse-grained region, which is derived from a coarse-grained potential. A thermostat takes care of thermally equilibrating the atomistic degrees of freedom reintroduced in the transition region. This simple setup turned out to be computationally robust; the calculation of structural and thermodynamics properties in AdResS compared with the calculations done in a subregion of equivalent size in a full atomistic simulation shows a highly satisfactory agreement for several test systems.[26,28–32] The computational robustness encouraged the investigation of the conceptual justification of the method on the basis of first principles of thermodynamics and statistical mechanics.[33,34] This analysis first led to the introduction of a thermodynamic force acting on the center of mass of the molecules in the hybrid region. The thermodynamic force is based on the principle of uniformizing, to the atomistic value, the chemical potential of each (space dependent) resolution[7] and then to the derivation of such a thermodynamic force from a more general thermodynamic principle, that is from the balance of grand potential for two interfaced open systems,[8]

$$\left[P_{atom} + \rho_o \int_\Delta F_{th}(x)dr\right] V = P_{CG}V, \quad (2)$$

where P_{atom} and P_{CG} are the pressure of the atomistic and coarse-grained regions, ρ_o is the target density of the reference full atomistic simulation, and V the volume of the simulation box. The explicit calculation of $F_{th}(x)$ is reported in Sec. III B 1. Based on such derivation, a step forward was done and AdResS was reformulated in terms of grand canonical formalism (GC-AdResS) where mathematical rigorous conditions were derived in order to assure that in the atomistic region, the system samples a grand canonical distribution. Such conditions, at the first order, has been shown to be equivalent to the use of the thermodynamic force.[10,11] Moreover, the coarse-grained model can be arbitrarily chosen without any reference to the atomistic model. In recent work,[12] we have embedded the scheme into the grand ensemble model of Bergmann and Lebowitz[35,36] and introduced a local thermostat acting only in the coarse-grained and hybrid region. Such a formalization allows one to define, with well founded physical arguments, the Hamiltonian of the atomistic (high resolution) region as the kinetic energy plus the interaction energy of the molecules in the atomistic region only; this implies that the interaction with molecules outside can be formally neglected. The definition of the Hamiltonian allows then to properly define the procedure for the calculation of equilibrium time correlation functions; moreover, for the case

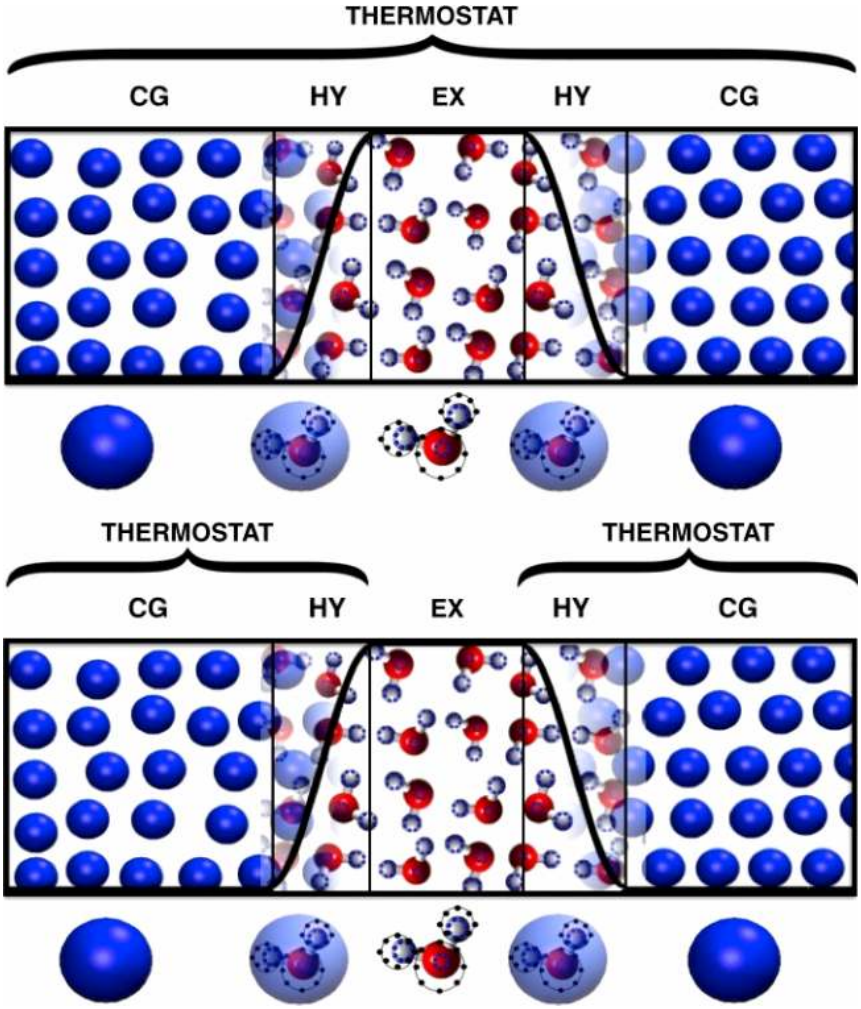

FIG. 1. Pictorial representation of the GC-AdResS scheme; CG indicates the coarse-grained region, HY the hybrid region where path-integral and coarse-grained forces are interpolated via a space-dependent, slowly varying, function $w(x)$ and EX (or PI) is the path-integral region (that is the region of interest). Top, the standard setup with the thermostat that acts globally on the whole system used in the calculation of static properties. Bottom, the "local" thermostat technique employed in this work in the calculation of dynamical properties.

of PI approach, this setup will provide a rigorous definition of the Hamiltonian of quantization. As it will be specified later on, there also exists a clear numerical argument that supports the definition of an accurate Hamiltonian in the high resolution region.

III. PIMD TECHNIQUES

The path integral formalism of Feynman applied to molecular simulation/dynamics of molecular systems is a well established approach and thus here, we will not report its formal derivation but only those aspects which are technically relevant for this specific study. A formal derivation and discussion of basic aspect of this approach can be found in Refs. 16 and 37, for example. The essential point of interest (in this paper) is the transformation, via path integral formalism, of a classical Hamiltonian of N distinguishable particles with phase space coordinate $(\mathbf{p},\mathbf{r})$, mass m_j (for the jth particle), and interaction potential in space $V(\mathbf{r}_1,\ldots,\mathbf{r}_N)$,

$$H = \sum_{j=1}^{N} \frac{\mathbf{p}_j^2}{2m_j} + V(\mathbf{r}_1,\ldots,\mathbf{r}_N), \tag{3}$$

into a quantized Hamiltonian which is formally equivalent to a Hamiltonian of classical polymer rings (atoms). The inter-atomic potential is distributed over the beads in such a way that each bead of a polymer ring interacts with the corresponding bead of another polymer ring. The intra-atomic interactions consist of harmonic potentials which couple each bead to the first neighbors in the chain. The fictitious dynamics of this polymeric liquid, with the spatial fluctuations/oscillations of the rings describing the quantum spatial delocalization of the atoms, allows for the calculation of quantum statistical properties of the atomic/molecular system. The quantized Hamiltonian takes the form

$$H_P = \sum_{i=1}^{P} \left[\sum_{j=1}^{N} \frac{[\mathbf{p}^{(i)}]_j^2}{2m'_j} + \sum_{j=1}^{N} \frac{1}{2} m_j \omega_P^2 (\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + \frac{1}{P} V(\mathbf{r}_1^i,\ldots,\mathbf{r}_N^i) \right], \tag{4}$$

where P is the number of beads of the polymer, $m'_j = \frac{Pm}{(2\pi\hbar)^2}$ and $\mathbf{p}^i$ are a fictitious mass and momentum, respectively, $\omega_P = \frac{\sqrt{P}}{\beta\hbar}$ ($\beta = 1/k_BT$), and $V(\mathbf{r}^i_1, \ldots, \mathbf{r}^i_N)$ is the potential that acts between same bead index i of two different particles. This setup allows to use molecular dynamics for the calculation of statistical properties. However, the direct use of the Hamiltonian above has shown to lead to a highly non-ergodic dynamics and suffers from poor sampling problems in the extended phase space of polymer rings,[3] since there are a wide range of frequencies present. The highest frequency limits the time step to be used in the simulation which causes the low frequency modes to be poorly sampled. Thus, either a very small time step should be used or very long runs should be performed, starting from different initial conditions. In order to circumvent the ergodicity problem, normal modes transformation is preferred.[16,38] The basic idea is to decouple the harmonic spring term, so that only a single harmonic frequency remains in the dynamics, and the time step for the simulation can be adjusted accordingly. The whole procedure is based on a transformation of coordinates to normal mode coordinates and thus to the use of an effective Hamiltonian,

$$H_P = \sum_{i=1}^{P}\left[\sum_{j=1}^{N}\frac{p_j^{(i)^2}}{2m_j^{(i)'}} + \sum_{j=1}^{N}\frac{1}{2}m_j^{(i)}\omega_P^2 (x')_j^{(i)^2} + \frac{1}{P}V\left(\mathbf{r}_1^{(i)}(x'_1), \ldots, \mathbf{r}_N^{(i)}(x'_N)\right)\right], \tag{5}$$

where $\frac{1}{P}U\left(r_1^{(i)}(x'_1), \ldots, r_N^{(i)}(x'_N)\right)$ is the potential that acts between same bead index i of two different particles in terms of the normal mode coordinates $x'_1, \ldots, x'_N$.

A. Choice of masses

In the standard PIMD,[39,40] the masses $m_j^{(i)'}$ are chosen such that all the internal modes have the same frequency and the sampling is efficient. Thus, the choice of mass is

$$m_j^{(i)'} = m_j\lambda_j^i, i = 2, \ldots, P,$$
$$m_j^{1'} = m_j,$$

where m_j is the physical mass and λ_j^i are the eigenvalues obtained by the normal mode transformation. This approach was used to calculate static properties and here we will use it, within GC-AdResS, for the same purpose. We will refer to this approach as **H1**. Craig and Manolopoulos[41] have developed RPMD, which has been successfully shown to calculate time correlation functions; the choices of the masses in RPMD are as follows:

$$m_j^{(i)'} = m_j. \tag{6}$$

In this work, we will employ this approach within GC-AdResS to calculate, in addition to static properties, time correlation functions; we will refer to it as **H2** approach. However, there exists an alternative formulation for RPMD.[19] The classical Hamiltonian for RPMD is

$$H_P = \sum_{i=1}^{P}\left[\sum_{j=1}^{N}\frac{[\mathbf{p}^{(i)}]_j^2}{2m_j} + \sum_{j=1}^{N}\frac{m_j}{2\beta_P^2\hbar^2}(\mathbf{r}_j^{(i)} - \mathbf{r}_j^{(i+1)})^2 + V(\mathbf{r}_1^i, \ldots, \mathbf{r}_N^i)\right], \tag{7}$$

where $\beta_P = \beta/P$, which effectively means that the simulation is performed at P times the original temperature. Moreover, the harmonic bead-bead interaction and the potential energy are scaled by P relative to Eq. (5). In Ref. 18, equivalence between different RPMD formalisms was shown. Due to the calculation of the thermodynamic force, for GC-AdResS simulations, this becomes an interesting technical aspect to investigate (see Sec. III B 1). We will refer to this approach as **H3** and verify its numerical robustness in GC-AdResS by comparing its results with the results obtained from **H1** and **H2**.

B. PIMD in GC-AdResS

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 than the dynamics of classical degrees of freedom; thus, the standard AdResS could be applied (technically) in the same way, with only one modification,[13–15]

$$F_{\alpha\beta} = w(X_\alpha)w(X_\alpha)F^{PI}_{\alpha\beta} + [1 - w(X_\alpha)w(X_\alpha)]F^{cm}_{\alpha\beta}, \tag{8}$$

where $F^{PI}_{\alpha\beta}$ is the force between beads of the rings representing the atoms of molecule α and molecule β. There exists another version of AdResS, recently developed in the group of Kremer,[42] based on a global Hamiltonian where the atomistic and coarse-grained potentials are interpolated instead of forces. The force-based and Hamiltonian-based approaches were shown to be numerically equivalent,[11,43] while conceptually both methods need strong assumptions when considering the physics of the entire system (see discussion in Refs. 12 and 44–46). In our force-based approach, from the conceptual point of view, the coupling between the polymer rings and the coarse-grained molecules cannot be rigorously expressed in a Hamiltonian form. However, calculations have shown that PIMD-AdResS was able to reproduce very well results obtained with full PIMD simulations. Since a rigorous Hamiltonian formalism is at the basis of the PIMD approach, the procedure of Refs. 13–15 was empirical and could be verified only *a posteriori*. The reason why the procedure was successful is that the coupling between the polymer rings and the coarse-grained molecules is negligible, in terms of energetic contribution, under the hypothesis that the path integral region and the coarse-grained region were large enough compared to the hybrid region. However, we have also numerically verified that even when all the three regions are relatively small and comparable in size, results are still satisfactory. The latest formalization of AdResS in GC-AdResS, reported in Sec. II, justifies why from a conceptual point of view the setup of PI-AdResS is robust. In fact, according to the model of Bergmann and Lebowitz,[12,35,36] for a simulation in a grand ensemble, one does not need to have an explicit coupling between the path integral region and the reservoir. The necessary and sufficient condition is the knowledge of the molecules' distribution in the reservoir. It follows that the interaction of the molecules of the path integral region with the rest of the system, while technically convenient and numerically efficient, from the conceptual (formal) point of view instead does not play a crucial role. Such an interaction plays only the technical role of a sort of "capping potential" which avoids that molecules entering the path integral region overlap in space. Moreover, the action of the thermodynamic force and of the thermostat in the hybrid region makes the stochastic coupling dominant (compared to the explicit hybrid interactions), which is the essence of any grand ensemble scheme. It follows that in GC-AdResS, the Hamiltonian to be considered for the path integral formalism is the Hamiltonian of the path integral region only, without any external additional term, i.e., the path integral region with its quantized Hamiltonian is embedded in a large reservoir with the proper grand canonical behaviour. It must be clarified that while the Bergmann-Lebowitz model provides an elegant and solid formal structure to the PI-AdResS, it is not strictly required to justify the existence of an accurate Hamiltonian in the PI region and thus the implementation of PIMD in AdResS. In fact, in the Appendix, we provide a numerical proof that, for the systems treated in this paper, the interaction energy between the PI region and the rest of the system is at least one order of magnitude smaller than the interaction energy of the molecules in the PI region. The accuracy and robustness of PI-AdResS (or PI-GC-AdResS) will be shown with the simulation of liquid water in Sec. IV. Finally, it must be clarified that for the current implementation of PIMD in GC-AdResS (in the GROMACS package), it is difficult to estimate the computational gain since the code architecture is not yet optimized. At this stage, we only want to show that the approach is satisfactory from a conceptual point of view. However, for very large systems with $P = 32$, the computational gain is around 1.7-2.0 compared to the full PIMD simulations. With further code modifications (e.g., removal of explicit degrees of freedom in the coarse-grained region, using multiple time steps) or with the implementation of PI-AdResS in a platform explicitly designed for PIMD simulations we estimated, for systems of the order of thousand molecules, a gain of at least a factor 4.0-5.0 compared to the full PIMD simulations.

1. Calculation of the thermodynamic force in PIMD

For an atomistic system, the thermodynamic force, $F_{th}(x)$, can be expressed as

$$F_{th}(x) = \frac{M}{\rho_o}\nabla P(x), \tag{9}$$

where M is the mass of the molecule and $P(x)$ is the pressure which characterizes different resolutions (for the initial configuration). $P(x)$ is approximated in terms of linear interpolation of molecular number density,

$$P(x) = P_{atom} + \frac{M}{\rho_o \kappa}\left[\rho_o - \rho(x)\right], \tag{10}$$

where κ is the compressibility and $\rho(x)$ is the density generated if the simulation runs without any thermodynamic force. The thermodynamic force is then obtained by an iterative procedure,

$$F^{th}_{k+1}(x) = F^{th}_{k}(x) - \frac{M_\alpha}{[\rho_o]^2 \kappa}\nabla \rho_k(x). \tag{11}$$

After each iteration, a density profile $\rho(x)$ is obtained due to the application of the thermodynamic force. The process converges when the density profile obtained is equal to the target density. At this point, the system is in thermodynamic equilibrium and the production run can start. The calculation of thermodynamic force in PIMD-GC-AdResS is essentially based on the same principle of balancing grand potential for interfaced open systems,

$$\left[P_{quantum} + \rho_o \int_\Delta F_{th}(r)dr\right] V = P_{CG} V, \tag{12}$$

where ρ_o is the target density of the reference full path-integral system. As for the classical case, $P(x)$ can be written as

$$P(x) = P_{quantum} + \frac{M_a}{\rho_o \kappa}\left[\rho_o - \rho(x)\right]. \tag{13}$$

While the above approach is highly efficient for classical simulations, for path integral simulations, it is cumbersome to run a PIMD-GC-AdResS simulation to calculate the thermodynamic force, before doing an actual production run, as the path-integral simulation is inherently very expensive. In order to make the scheme efficient, we have devised a strategy to calculate the thermodynamic force which requires least computation. As discussed in Sec. III A, we will show how the thermodynamic force is calculated for different Hamiltonian

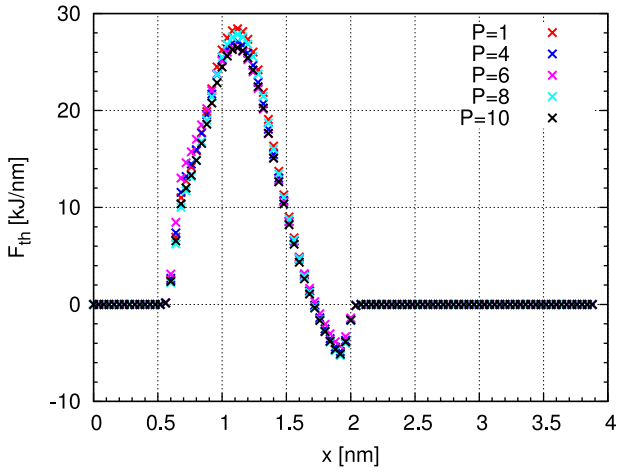

FIG. 2. Thermodynamic force calculated in AdResS simulation using **H1** (**H2**) approach. The force is calculated for different number of polymer ring beads. It does not change as the number of beads is varied.

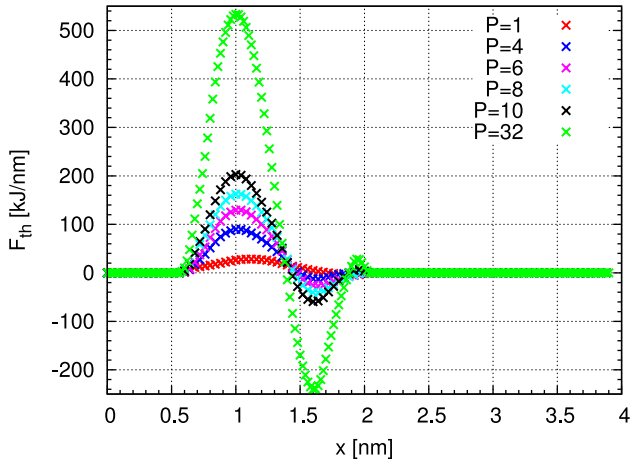

FIG. 3. Thermodynamic force calculated in AdResS simulation using **H3** approach. The force is calculated for different number of polymer ring beads. The thermodynamic force for $P = 32$ is then extrapolated by using space-dependent scaling factors calculated using thermodynamic force for $P = 1, 4, 6, 8$, and 10.

approaches. In case of **H1** and **H2**, where the temperature of the system is just the normal temperature, we calculated the thermodynamic force for path-integral systems with varying Trotter number $P = 1, 4, 6, 8$, and 10 ($P = 1$ represents the classical limit). Since the thermodynamic force takes care of a thermodynamic equilibration and since the thermodynamic conditions (thermodynamic state point) of a classical and a quantum system are the same, we expect that the thermodynamic force calculated in the classical case ($P = 1$) would be sufficient to provide thermodynamic equilibrium in simulations where $P = 32$ is used. In fact, we found that the thermodynamic force was same in all the cases. Fig. 2 shows the thermodynamic force calculated for a water system, with different number of ring polymer beads in each case. Using this argument, we used this thermodynamic force in the actual production run with $P = 32$. We found that the density of water molecules in the full quantum subregion and the transition region is equal to the reference density of the water system at the same thermodynamic conditions. Thus, in **H1** and **H2** approaches, if the quantum effects on the pressure of the system are not large, we can directly use the thermodynamic force calculated from the classical simulation.

In **H3** approach, the situation is more complex, as the effective temperature of simulation changes if the number of beads is changed, thus the (numerical) thermodynamic state point changes. In this case, there would be no other choice but to run a full PIMD-GC-AdResS simulation with $P = 32$ and calculate the thermodynamic force. However, we avoided such an expensive calculation and instead calculated the thermodynamic force for system with different number of beads $P = 1, 4, 6, 8$, and 10 at temperatures $T = 298 \times P$ and extrapolated thermodynamic force for $P = 32$, using space dependent factors calculated from thermodynamic force for smaller values of P (Fig. 3). Next, we used this thermodynamic force for production run with $P = 32$ and found that the density of water molecules in the full-PI subregion is same as the target density, while the density in transition region differs at worst by 3%.

2. Equilibrium time correlation functions: Theoretical and computational aspects

The technique of RPMD (**H2**) focuses on the Kubo-transformed correlation functions.[47,48] The Kubo-transformed correlation function of the operators $\hat{A}$ and $\hat{B}$ is defined by[41]

$$K_{AB}(t) = \frac{1}{\beta Z}\int_0^\beta d\lambda \left[e^{-(\beta-\lambda)\hat{\mathcal{H}}}\hat{A}e^{-\lambda\hat{\mathcal{H}}}e^{i\hat{\mathcal{H}}t/\hbar}\hat{B}e^{-i\hat{\mathcal{H}}t/\hbar}\right], \tag{14}$$

where Z is the canonical partition function,

$$Z = tr\left[e^{-\beta\hat{\mathcal{H}}}\right]. \tag{15}$$

The RPMD method approximates the Kubo-transformed correlation functions by using the classical ring-polymer trajectories generated by the dynamics produced by the Hamiltonian in Eq. (7). The RPMD approximation is given by[49]

$$\tilde{c}_{AB}(t) \approx \frac{1}{(2\pi\hbar)^{9PN}Z_P}\int\int d^P p_0 d^P r_0 e^{-\beta_P H_P(p_0,r_0)} \times \frac{1}{N}\sum_{i=1}^{N} A_P^i(r_0)B_P^i(r_t), \tag{16}$$

where Z_P is the canonical partition function, and r_t indicates the time evolution at time t of the positions. The functions $A_P(r_o)$ and $B_P(r_t)$ are calculated by taking the average over the beads of the ring polymer,

$$A_P(r) = \frac{1}{P}\sum_{j=1}^{P} A(r), B_P(r) = \frac{1}{P}\sum_{j=1}^{P} B(r). \tag{17}$$

For the calculations in GC-AdResS, the above equation needs to be written in the formalism of the grand canonical ensemble,

$$\tilde{c}_{AB}(t) \approx \frac{1}{Z_P^{GC}}\sum_N \frac{1}{(2\pi\hbar)^{9PN}N!}\iint d^P p_0(N) d^P r_0(N) e^{-\beta_P H_P(N)(p_0(N),r_0(N))-\mu N} \times \frac{1}{N'}\sum_{i=1}^{N'} A_P^i(r_0(N))B_P^i(r_t(r_0(N))), \tag{18}$$

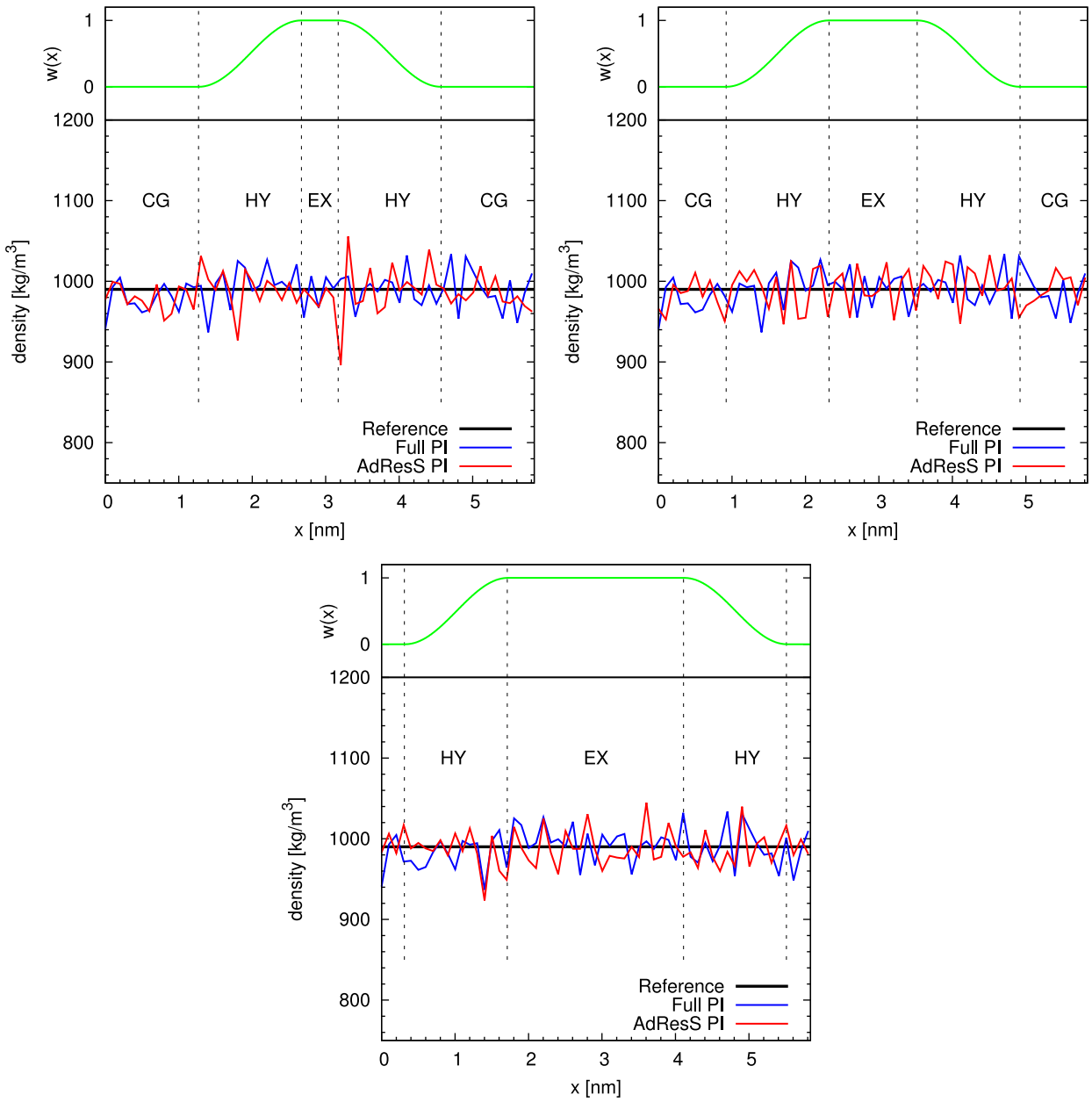

FIG. 4. Molecular number density calculated with GC-AdResS for different sizes of quantum subregion. Results are compared with the density obtained in a full path integral simulation.

where μ is the chemical potential and N' is the number of molecules at time "0" that remain correlated at time "t" (that is, molecules which remain in the path integral region for the whole time within the time window considered); $Z_P^{GC} = \sum_N e^{\beta\mu N} Z_P$ is the grand-canonical partition function and $H_P(N)$ is the Hamiltonian of the (open) path integral region with N instantaneous number of molecules. It must be noticed that the *a priori* knowledge of μ is not required; actually, in GC-AdResS, μ is automatically calculated by the equilibration procedure of the thermodynamic force (see also Ref. 11). From the technical point of view, we have used the same calculation procedure as that of Ref. 12, where equilibrium time correlation functions were calculated in the open subsystems using classical molecular dynamics. Such a principle is based on the definition of reservoir in the Bergmann-Lebowitz model, which implies that when a molecule leaves the system and enters the reservoir, it looses its microscopic identity and thus the corresponding correlations; thus, if a molecule which is present at time t_0 disappears from the system at time t (i.e., moves into the reservoir), then the contribution of this molecule, outside the time window $[0,t]$, to the correlation function shall not be considered. In our previous work, we have shown that such a principle is physically consistent on the basis of results of molecular simulations. Since all the beads in a ring-polymer are treated as dynamical variables,[18] there are no thermostats used in full RPMD simulations. Thus, the simulations are performed under NVE conditions, with either starting configurations generated from massively thermostated PIMD simulations[50] or re-sampling of momenta from Maxwell-Boltzmann distribution after every few picoseconds.[51] In order to keep the dynamics of the beads Newtonian in the path-integral subregion of GC-AdResS, we use a

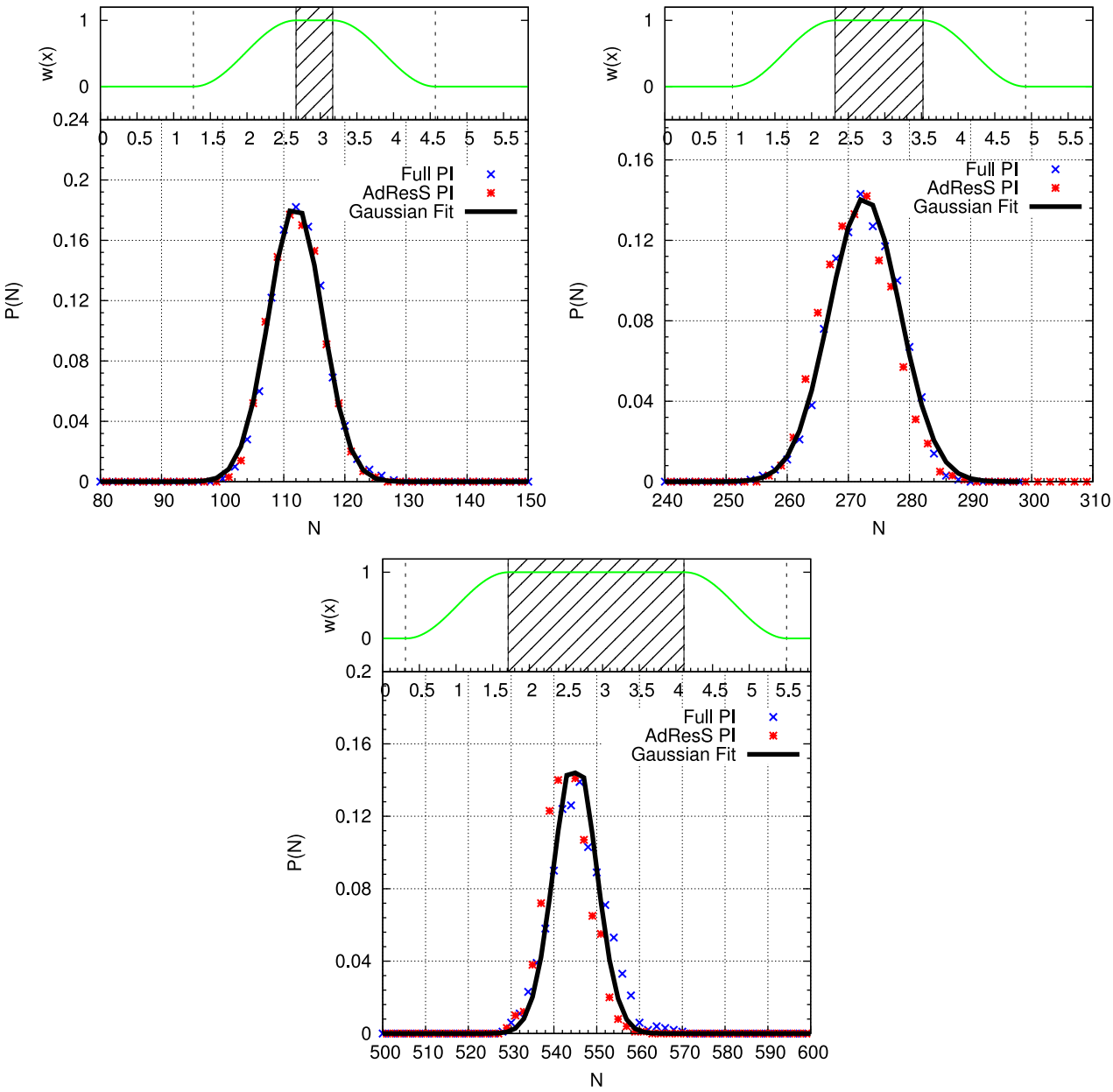

FIG. 5. Particle number probability distribution of GC-AdResS compared with the equivalent full path integral subsystem for different sizes of the quantum subregion. The shape of both curves is a Gaussian (reference black continuous curve) in all the three different simulations. The top part of the figure indicates the extension of the PI region (compared to the rest of the system) where the function is calculated; this representation is equivalent in all subsequent figures.

"local-thermostat" procedure,[12,52] where the thermostat is applied only in the coarse-grained and hybrid region, while the explicit path-integral region is thermostat-free. This ensures that the molecules which are present in the path-integral subregion are not subject to any perturbation due to the direct action of the thermostat.

IV. RESULTS

In this section, we report results about the simulation of liquid water at room conditions. The quantum model for liquid water used in this work is q-SPC/FW.[58] It was shown that the thermodynamic and dynamical properties calculated using this water model agree quite well with the experimental data. The section is divided in two parts, the first where static results (molecular number density across the system, radial distribution functions, and probability distribution of the molecules) are reported, and the second where several equilibrium time correlation functions are calculated. Few further points must be mentioned as clarification to this study. The total volume of the PIMD-GC-AdResS box is the same in all simulations, while three different sizes of the region at PI resolution are used and the dimension of the transition region is always kept the same. The smallest size of the PI region represents the limiting case of a statistically relevant number of molecules treated with PI resolution. The largest size instead represents the limiting case of a reservoir (hybrid plus coarse-grained region) which is relatively small and thus, it may be expected to not fulfill the conceptual requirement of being statistically large enough. We will show that even in these two limiting cases, the method

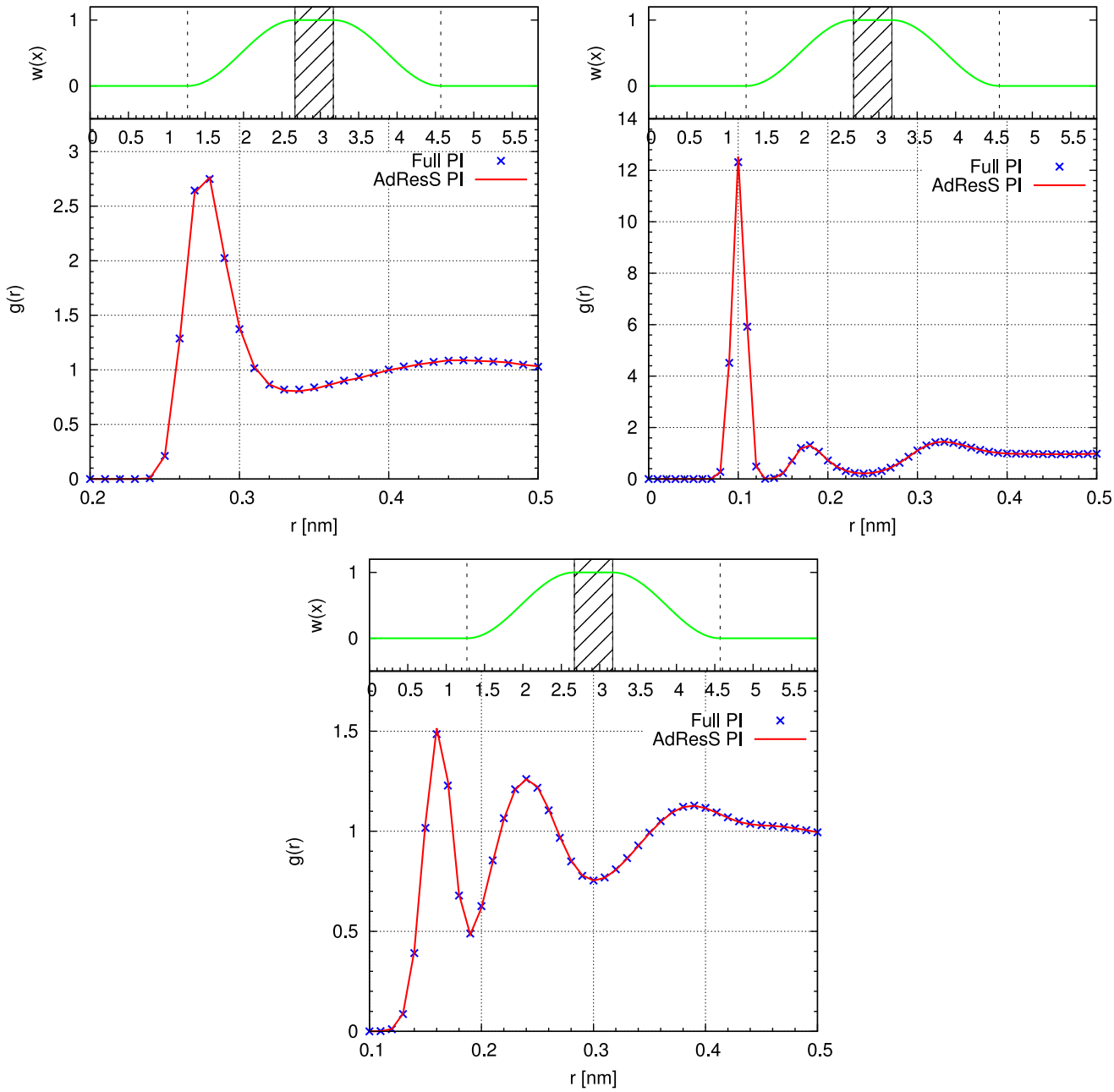

FIG. 6. From left to right: (Bead-bead) oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen partial radial distribution functions calculated with path integral AdResS. Such functions are compared with the results obtained for an equivalent subsystem ($EX = 0.5$ nm) in a full path integral simulation.

is computationally and conceptually robust. A second point to take into account is that we compare the results of GC-AdResS for the PI region with the results obtained in a subsystem of a full PI simulations, such a subsystem is of the same size as in the GC-AdResS simulation. The subsystem of a large full PI simulation box is a natural grand canonical ensemble; thus, if our subsystem of AdResS reproduces the results of a full PI subsystem, then we can be rather confident that the PI region in AdResS samples the Grand Canonical distribution sufficiently well. From the physical point of view, it should be clarified that the functions calculated in a subsystem must be considered local in space and time if compared to calculation done over the whole simulation box of the full PI simulation. Once again, as the subsystem size increases, the functions go to the value obtained in a full PI simulation when the full box is considered (for physical consistency, see checks in Ref. 12). Technical details of the simulation are reported in the Appendix.

A. Static properties

We use the **H1** and **H2** PIMD approaches (**H1**-GC-AdResS and **H2**-GC-AdResS, respectively, for the GC-AdResS simulation), Fig. 4 shows molecular number density. In all three cases, the agreement is highly satisfactory, the largest deviation is found for the case with PI region of 0.5 nm and is below 5%. This is the basic test to show equilibration and thermodynamic consistency; moreover, following the mathematical formalization of Ref. 10, is the first order necessary condition in order to have the correct grand canonical distribution in the PI region. A further confirmation of the

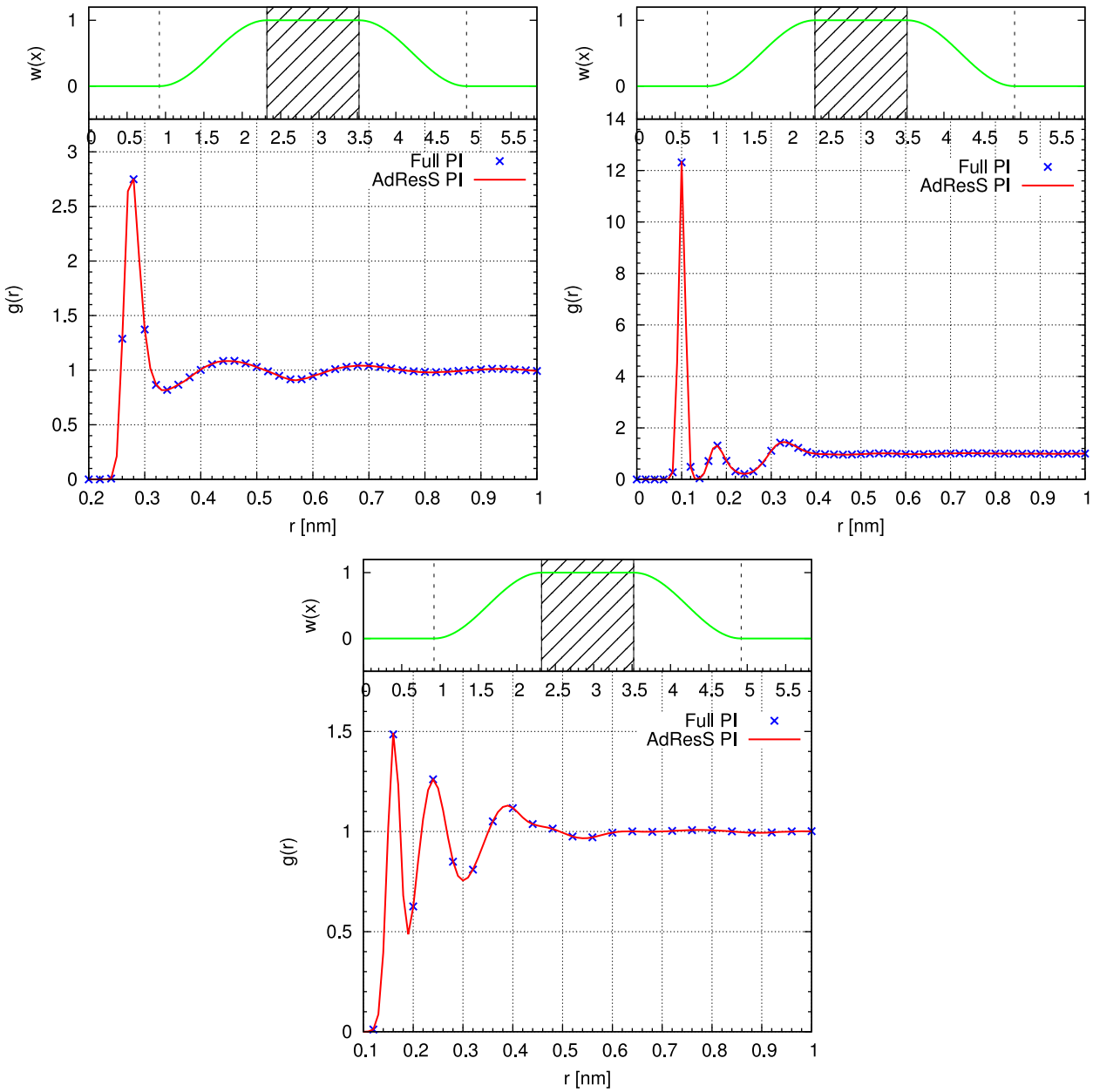

FIG. 7. From left to right: (Bead-bead) oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen partial radial distribution functions calculated with path integral AdResS. Such functions are compared with the results obtained for an equivalent subsystem ($EX = 1.2$ nm) in a full path integral simulation.

fact that the method samples the phase space of a subsystem in a sufficiently correct way is represented by Fig. 5. The figure shows the particle number probability distribution in quantum subregion of AdResS and an equivalent subregion in full path integral simulation. It can be seen that also in this case, the results are highly satisfactory and the shape of two curves is a Gaussian, as one should expect. There is a systematic shift by (at most) two-three particles between the full path integral results and the adaptive resolution scheme for particle number distributions; such a shift seems to correlate with the width of the distribution (i.e., extension of the PI region). An explanation of the effect is that in AdResS, the average particle density is not perfectly flat across the box and for larger PI regions, the frequency in space of small density fluctuations is larger than for smaller PI regions. This is due to the fact that we have used an empirical approach to evaluate the thermodynamic force instead of the standard one, as explained before. However, the discrepancy is numerically negligible. The g(r) is an important structural quantity that represents a two-body correlation function and thus a higher order than the molecular density of the ensemble many-body distribution; moreover, it differs considerably when quantum models of water are used, in particular correlation functions involving hydrogen atoms.[53] We calculated the local bead-bead g(r)'s in the quantum subregion in GC-AdResS and compared them with the bead-bead g(r)'s in an equivalent subregion in the full path-integral simulation. Figs. 6–8 show that the results from GC-AdResS agree with the results from full PI simulation in a highly satisfactory way.

We have also verified, for the most relevant case ($EX = 1.2$), that also **H3** approach gives satisfactory results for the static properties when employed in GC-AdResS; results are

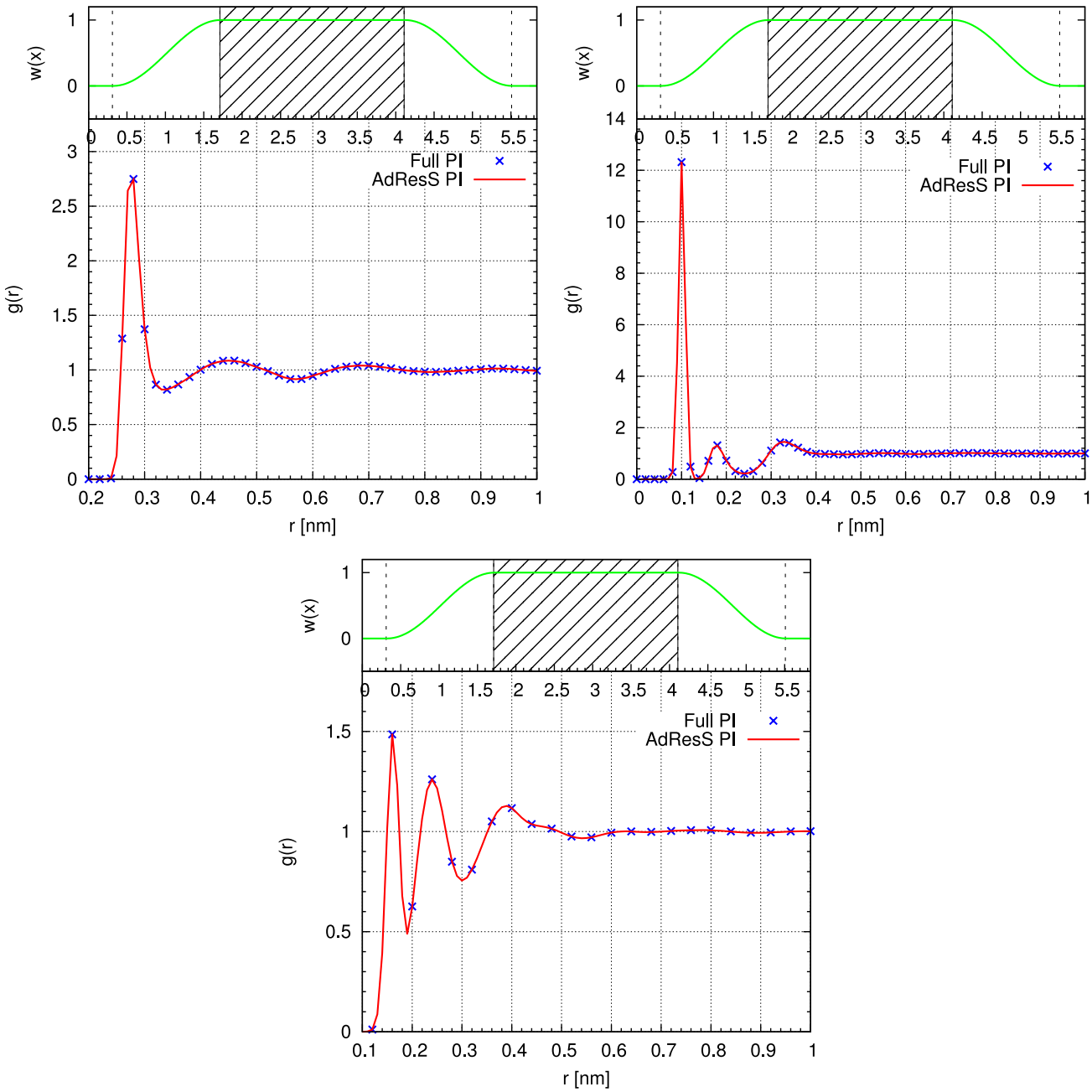

FIG. 8. From left to right: (Bead-bead) oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen partial radial distribution functions calculated with path integral AdResS. Such functions are compared with the results obtained for an equivalent subsystem ($EX = 2.4$ nm) in a full path integral simulation.

reported in Fig. 9. Due to the more empirical calculation of the thermodynamic force in **H3**, results are not as accurate as those of **H1** and **H2**; the density in the hybrid region differs by around 3%, which is anyway numerically negligible (however, the difference must be reported). However, the number probability distribution and bead-bead g(r)'s agree quite well in AdResS and full path-integral simulations. This leads to the conclusion that also results obtained with **H3** are highly satisfactory.

This section essentially shows the ability of PI-GC-AdResS with all three PIMD techniques to sample basic (but highly relevant) static properties of a grand canonical ensemble. In order to prove that a more elaborated sampling is also satisfactorily made by the method, we report in Sec. IV B the calculation of equilibrium time correlation functions.

B. Dynamic properties

We report results for the velocity-velocity autocorrelation function, for the first and second order orientational (molecular dipole) correlation functions,[54,55] and for the reactive flux correlation function for hydrogen bond dynamics.[56,57] This latter in specific situations may strongly diverge from the classical case, and thus, it may be a quantity of relevance for this work. Moreover, the fact that PI-GC-AdResS reproduces the behaviour of a full PI simulation is of high technical relevance in perspective (e.g., study of solvation of molecules). The explicit formulas used for the functions calculated here are given in Ref. 61. All results shown in this section are highly satisfactory, either when **H2** is used or **H3** is used. Thus, the

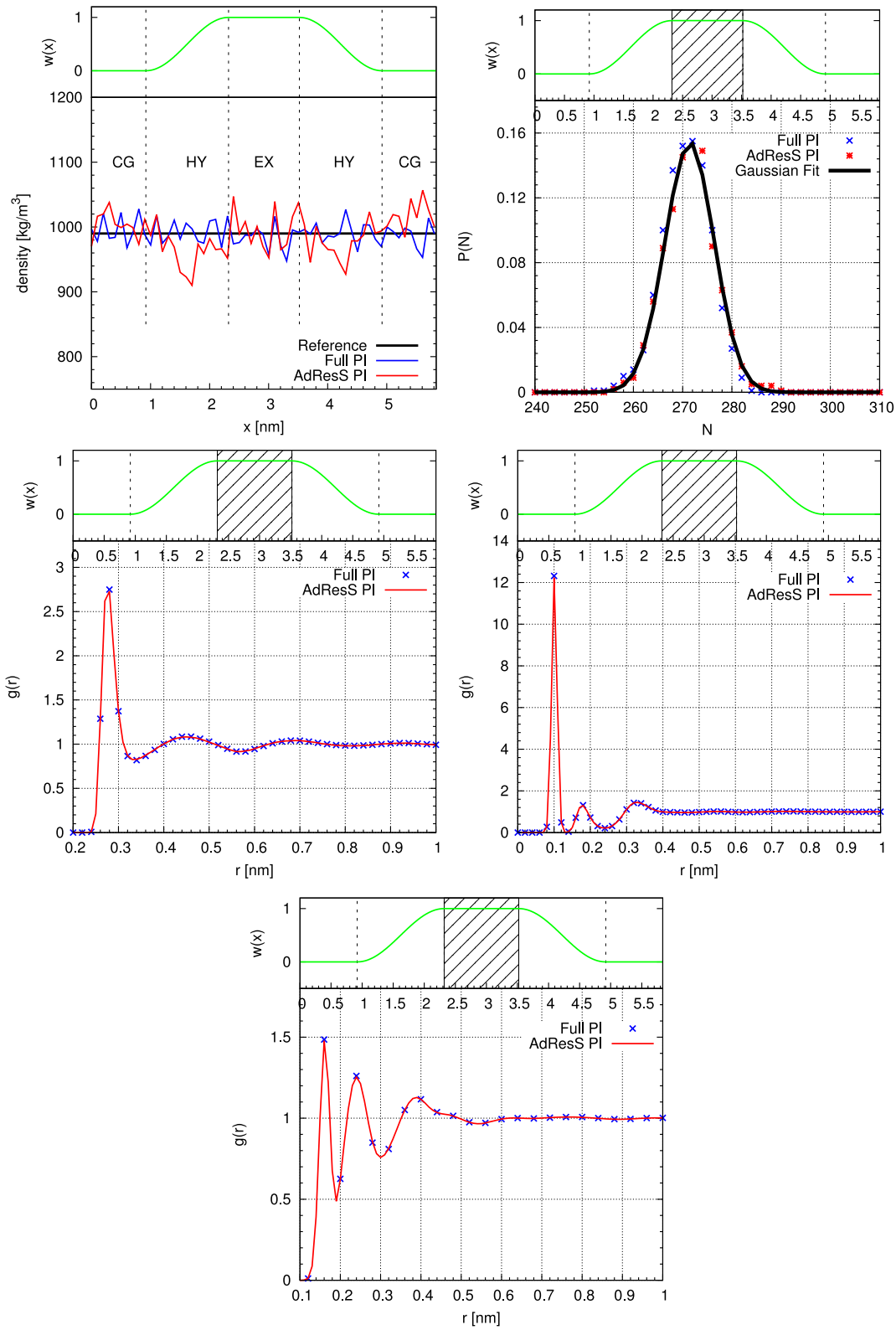

FIG. 9. From left to right (top): Particle number density, Particle number probability distribution of GC-AdResS obtained using the **H3** approach. From left to right (bottom): (Bead-bead) oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen partial radial distribution functions calculated with path integral AdResS using the **H3** approach. Such functions are compared with the results obtained for an equivalent subsystem ($EX = 1.2$ nm) in a full path integral simulation.

PI-GC-AdResS can be certainly considered a robust computational method for the calculation of quantum-based static and dynamic properties of liquid water and as a consequence for simpler systems and for systems where water plays a major role (at least).

1. Equilibrium time correlation functions

Figures 10–12 show the three correlation functions calculated in the quantum subregion in GC-AdResS and in an equivalent subregion in RPMD simulation, where the explicit region

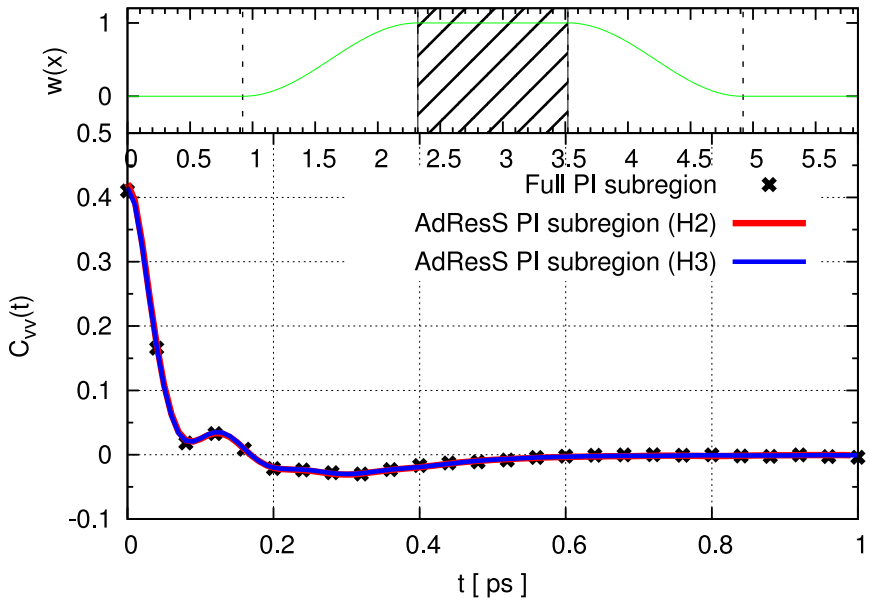

FIG. 10. Kubo-transformed velocity autocorrelation function for q-SPC/FW water model calculated in the quantum subregion of GC-AdResS and in an equivalent subregion in RPMD simulation.

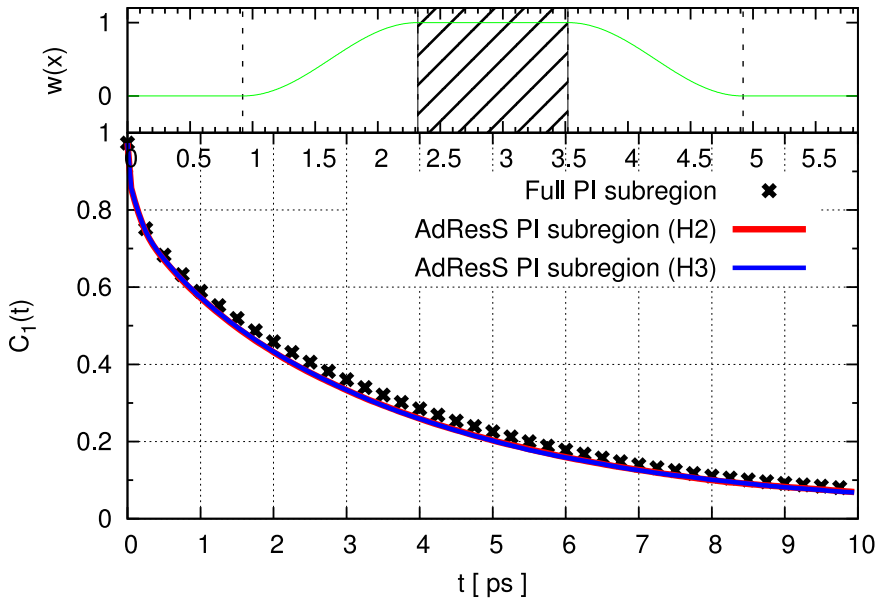

FIG. 11. Kubo-transformed first order orientational correlation function for q-SPC/FW water model calculated in the quantum subregion of GC-AdResS and in an equivalent subregion in RPMD simulation. Dipole moment axis is chosen as the inertial axis of molecule.

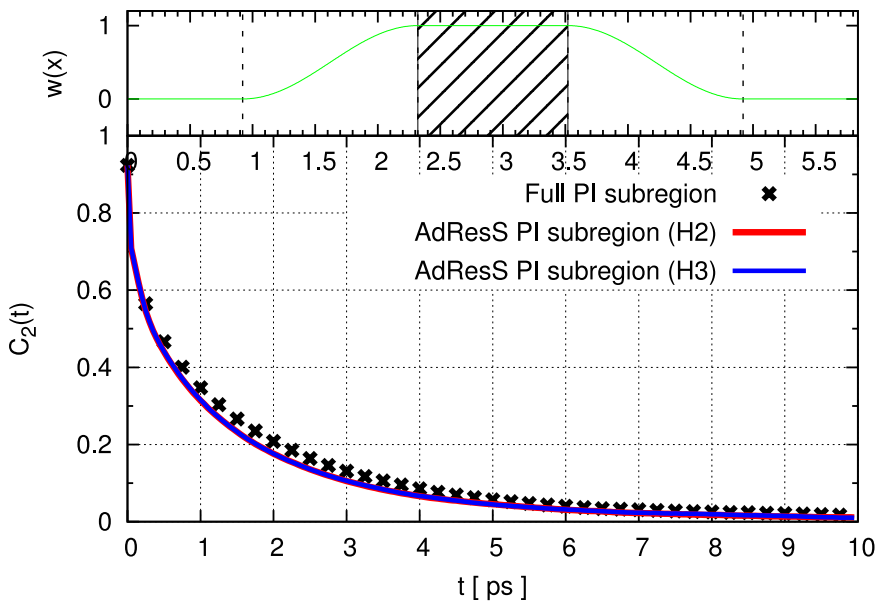

FIG. 12. Kubo-transformed second order orientational correlation function or q-SPC/FW water model calculated in the quantum subregion of GC-AdResS and in an equivalent subregion in RPMD simulation. Dipole moment axis is chosen as the inertial axis of molecule.

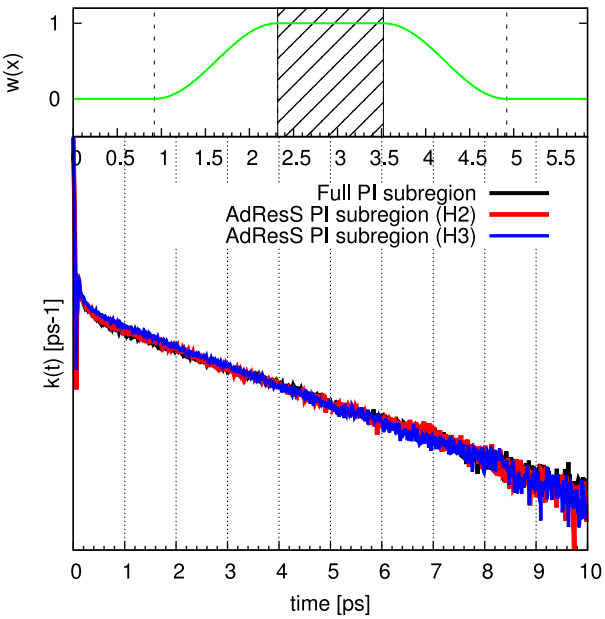

FIG. 13. The rate function k(t) (semilogarithmic plot) for q-SPC/FW water model calculated in the quantum subregion of GC-AdResS and in an equivalent subregion in RPMD simulation.

is 1.2 nm. All the correlation functions are calculated using **H2** approach and **H3** approach; results confirm the consistency of the two methods in GC-AdResS. As stated before, these are the local time correlation functions, calculated in the specific region of interest, and could differ from the global time correlation functions, calculated over the whole system. However, it was shown in Ref. 12 that as the size of the explicit region increases, the local correlation functions converge to the global correlation functions.

2. Dynamics of hydrogen bonding

In order to investigate the dynamics of hydrogen bond formation and breaking using RPMD simulations, we calculate the hydrogen bond population fluctuations in time, which are characterized by the correlation function,

$$c(t) = \langle h(0)h(t)\rangle/\langle h\rangle, \tag{19}$$

where $h(t)$ is the hydrogen bond population operator, which has a value 1, when a particular pair is bonded, and zero otherwise. One can then calculate the rate of relaxation as

$$k(t) = -dc/dt, \tag{20}$$

where $k(t)$ is the average rate of change of hydrogen-bond population for those trajectories where the bond is broken at a time t later. Two water molecules are treated as hydrogen bonded, if the distance between the center of two oxygen rings is less than 0.35 nm and, simultaneously, the angle between the axis defined by the center of two oxygen ring polymers and the center of one of the oxygen-hydrogen rings is less than 30°. Fig. 13 shows $k(t)$ calculated in the quantum subregion of AdResS and an equivalent subregion in RPMD simulation.

V. CONCLUSION

We have performed simulations of liquid water at room conditions using PIMD in three different technical approaches. Each of these approaches was embedded in GC-AdResS so that a PIMD for open systems in contact with a generic reservoir is realized. The results regarding static and dynamic quantities are highly satisfactory and qualify PI-GC-AdResS as a robust method for simulations of systems which currently are prohibitive for full PIMD simulations; for example, the already mentioned solvation problem. One can define a high resolution region at PI resolution around the solute and surround the solvation region with a reservoir as that constructed in GC-AdResS. The static and dynamic properties of the hydrogen bonding network can be analyzed and, by comparing results with those of classical systems, one may conclude about the importance of quantum effects due to hydrogen spatial delocalization. This approach can introduce not only a technical innovation regarding the computational efficiency but, by varying the size of the high resolution region, could also be used as a tool of analysis to identify the essential degrees of freedom required by a certain physical process. In this perspective, here we have shown that PI-GC-AdResS is a robust method for linking the microscopic to macroscopic scale in a truly multiscale fashion.

ACKNOWLEDGMENTS

We thank Professor Thomas Markland for a critical reading of the manuscript and for useful suggestions. We thank Christoph Junghans, Han Wang, and Adolfo Poma for many discussions during the formulation of the theory and the development of the code. We also thank Sebastian Fritsch for providing the path integral scripts. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) partially with the Heisenberg grant (Grant No. DE 1140/5-2) and partially with the Grant No. CRC 1114 provided to L.D.S. The DFG grant (Grant No. DE 1140/7-1) associated to the Heisenberg grant for A.G. is also acknowledged. Calculations were performed using the computational resources of the North-German Supercomputing Alliance (HLRN), Project No. **bec00127**.

APPENDIX: TECHNICAL DETAILS

1. Energetic contribution of the coupling term

The ith molecule (at position, $\boldsymbol{r}_i$) in the EX (PI) region is characterized by $w(\boldsymbol{r}_i) = 1$. It follows that the force acting on the ith molecule can be separated in two parts: (i) the force generated by the interaction of molecule i with molecules of the EX region,

$$\boldsymbol{F}_{i,j} = \boldsymbol{F}^{PI}_{i,j}, \forall j \in EX \tag{A1}$$

and (ii) the force generated by the interaction with molecules in the rest of the system,

$$\boldsymbol{F}_{i,j} = w(\boldsymbol{r}_j)\boldsymbol{F}^{PI}_{i,j} + [1 - w(\boldsymbol{r}_j)]\boldsymbol{F}^{CG}_{i,j}, \forall j \in HY + CG. \tag{A2}$$

From Eq. (A1), it follows

$$\boldsymbol{F}_i = \sum_{j\neq i} \boldsymbol{F}^{PI}_{i,j} = \sum_{j\neq i} \nabla_j U^{ij}_{PI}, \tag{A3}$$

where ∇_i is the gradient with respect to molecule i and U_{PI}^{ij} is a compact form to indicate the proper bead-bead interaction of atoms of molecule i with those of molecule j. Eq. (A2) represents instead the coupling force between molecules of $HY + CG$ region and molecule i, that is an external force. At this point, we argue that the non-integrable part of the dynamics in the HY region is a numerically negligible boundary effect. In fact, Eq. (A2) can be rewritten as

$$\begin{aligned} \boldsymbol{F}_i &= \sum_{j \in HY+CG} [w(\boldsymbol{r}_j)\boldsymbol{F}_{i,j}^{PI} + [1 - w(\boldsymbol{r}_j)]\boldsymbol{F}_{i,j}^{CG}] \\ &= \sum_{j \in HY+CG} [w(\boldsymbol{r}_j)\nabla_i U_{PI}^{ij} + [1 - w(\boldsymbol{r}_j)]\nabla_i U_{CG}]. \end{aligned} \quad \text{(A4)}$$

It follows that the energy of the ith molecule at a certain time t associated with the force of Eq. (A4) is given by

$$W_{PI-Res}^{i}(t) = \sum_{j \in HY+CG} [w(\boldsymbol{r}_j)U_{PI}^{ij} + [1 - w(\boldsymbol{r}_j)]U_{CG}^{ij}], \quad \text{(A5)}$$

where the $Res = HY + CG$. The total energy of coupling at time t is then defined as

$$W_{PI-Res}(t) = \sum_{i \in PI} W_{PI-Res}^{i}(t). \quad \text{(A6)}$$

In order to understand whether or not the quantity of Eq. (A6) is numerically negligible, one should compare it to the amount of energy, W_{PI-PI}, corresponding to the interaction between molecules of the PI region only: $W_{PI-PI}(t) = \sum_{i<j} U_{PI}^{ij}; i, j \in PI$. If

$$\frac{|W_{PI-PI}(t)| - |W_{PI-Res}(t)|}{|W_{PI-PI}(t)|} \approx 1; \forall t, \quad \text{(A7)}$$

then it seems reasonable to approximate the total energy of the PI region by the Hamiltonian of the PI region; thus, the Hamiltonian formalism is numerically justified in PI-AdResS. Fig. 14 shows that the difference in energy is at least of one order of magnitude and that condition (A7) holds in all simulations we have presented in this work. Moreover, it should be noticed that on purpose, we have performed simulations where the technical conditions are not optimal (the size of each region of the system is much smaller than the size prescribed by the theory); thus, Eq. (A7) would certainly hold in simulations with standard technical conditions.

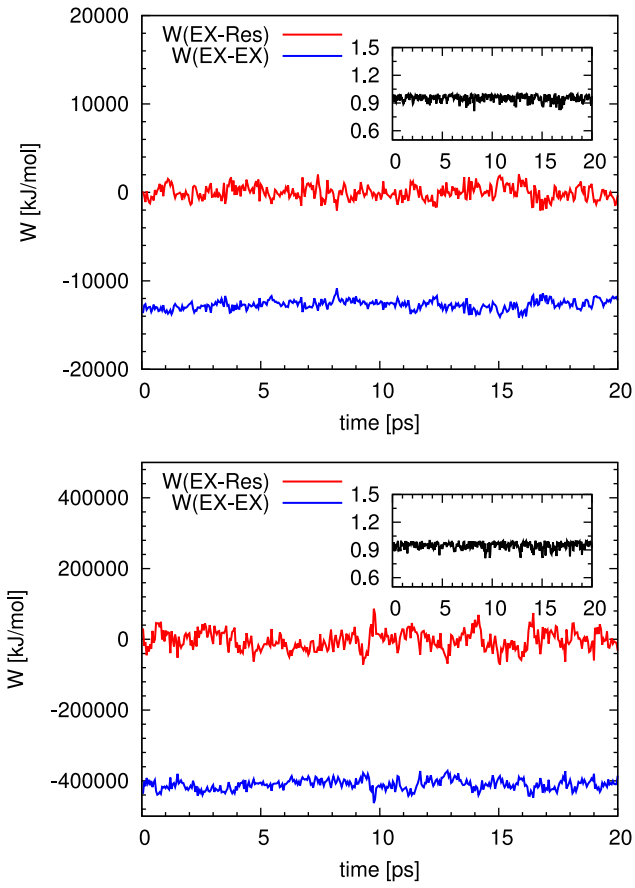

FIG. 14. Main figure: $W_{PI-PI}(t)$ compared to $W_{PI-Res}(t)$. Inset: The relative amount of the interaction between the PI region and the rest of the system along the trajectory: $\frac{|W_{PI-PI}(t)|-|W_{PI-Res}(t)|}{|W_{PI-PI}(t)|}$; the contribution is, at most, of 10%. Calculations are done within the **H1** and **H2** approaches (top) and **H3** (bottom).

2. Simulation setup

a. Static properties

All path integral simulations are performed by home-modified GROMACS,[59] and the thermodynamic force in GC-AdResS simulations is calculated using VOTCA.[60] The number of water molecules in system is 1320, and the box dimensions are $5.8 \times 2.6 \times 2.6$ nm^3, corresponding to a density 990 kg m^{-3}. In AdResS simulations, the resolution of the molecules changes along x-axis, as depicted in Figure 1. Three different AdResS simulations are performed, each with a different size of quantum subregion. The different sizes of the quantum subregion treated in this work are $0.5 \times 2.6 \times 2.6$ nm^3, $1.2 \times 2.6 \times 2.6$ nm^3, and $2.4 \times 2.6 \times 2.6$ nm^3. The transition region, which has dimensions $2.8 \times 2.6 \times 2.6$ nm^3, is fixed in all the three cases. The remaining system contains coarse-grained particles, which interact via generic WCA (Weeks-Chandler-Andersen) potential of the form

$$U(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon, r \leq 2^{1/6}\sigma. \quad \text{(A8)}$$

The parameters σ and ϵ in the current simulations are 0.30 nm and 0.65 kJ/mol, respectively. Thirty two ring polymer beads are used in all the simulations, which is sufficient to obtain the converged results for both static and dynamical properties. Reaction field method is used to compute the electrostatic properties with dielectric constant for water equal to 80. The cutoff for both van der Waals and electrostatic interactions is 1.2 nm. All the static properties are computed from 250 ps long trajectories. The simulations using **H1** and **H2** formalisms are performed at 298 K, while the simulations using **H3** formalism are performed at 9536 K. The time step used in all the simulations is 0.1 fs. In the calculation of the thermodynamic force, a single iteration consists of a 200 ps long trajectory which is used to compute the density profile. A total of 20 such iterations is sufficient to obtain a flat density profile and a converged thermodynamic force.

b. Dynamic properties

The system details are kept same as in Subsection 2 a of the Appendix. A 200 ps long PIMD simulation is performed and

along the trajectory, configurations are taken after every 8 ps to perform RPMD simulations. Thus, a total of 25 trajectories each of length 25 ps is generated. For the first 5 ps, we keep the thermostat switched on, in order to adjust the velocities as masses are different in PIMD and RPMD methods. After this initial equilibration run, the thermostat is switched off, and the NVE trajectories generated are used to compute various time correlation functions. We use the same strategy for AdResS simulations, where a 200 ps long fully thermostated GC-AdResS PIMD simulation is performed, and 25 initial configurations are taken along this trajectory to perform GC-AdResS RPMD simulations. For the first 5 ps, the thermostat acts in the explicit as well as the hybrid and coarse-grained regions. After the short equilibration run, the thermostat is switched off in the explicit region, while the hybrid and coarse-grained regions are kept under the action of the thermostat. The dynamic properties are calculated in the explicit region in the last 20 ps, i.e., excluding the equilibration run. The velocity auto-correlation function is calculated for 1 ps, while the orientational correlation functions and reactive flux correlation functions for hydrogen bond dynamics are calculated for 10 ps in one single trajectory and then averaged over all the trajectories.

3. Thermostat issue

It is well known that massive thermosetting is needed in the path integral simulations, as the forces arising due to the high frequencies in the polymer ring and the forces due to the potential $U(x)$ are weakly coupled. Tuckerman *et al.*[16] coupled each normal mode variable to separate Nose-Hoover chains, thereby ensuring proper ergodic sampling of the phase space. Manolopoulos *et al.*[19] developed specific Langevin equations for thermostat that are tuned to sample all the internal modes of the ring polymer quite efficiently. However, in this work, we chose the standard Langevin equations of thermostat with time scale 0.1 ps, which is strong enough for sampling the phase space effectively, though it may not be the most efficient choice. The reason is that in the initial stage of validating GC AdResS for path integral simulations, we need to show that the properties obtained in the full PI simulations are reproduced exactly in AdResS. Since we use the same thermostat in both the simulations, there should not be any discrepancy arising due to the thermostat. However, the comparison of static properties calculated in our reference PIMD simulation with those available in the literature (referring to the approaches above) is highly satisfactory.

[1]R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, Inc., 1965).

[2]H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets* (World Scientific Publishing Co., 2009).

[3]M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, New York, 2010).

[4]P. Ball, Chem. Rev. **108**, 74 (2008).

[5]M. Praprotnik, L. Delle Site, and K. Kremer, J. Chem. Phys. **123**, 224106 (2005).

[6]M. Praprotnik, L. Delle Site, and K. Kremer, Annu. Rev. Phys. Chem. **59**, 545 (2008).

[7]S. Poblete, M. Praprotnik, K. Kremer, and L. Delle Site, J. Chem. Phys. **132**, 114101 (2010).

[8]S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, and K. Kremer, Phys. Rev. Lett. **108**, 170602 (2012).

[9]H. Wang, C. Schütte, and L. Delle Site, J. Chem. Theory Comput. **8**, 2878 (2012).

[10]H. Wang, C. Hartmann, C. Schütte, and L. Delle Site, Phys. Rev. X **3**, 011018 (2013).

[11]A. Agarwal, H. Wang, C. Schütte, and L. Delle Site, J. Chem. Phys. **141**, 034102 (2014).

[12]A. Agarwal, J. Zhu, C. Hartmann, H. Wang, and L. Delle Site, New J. Phys. **17**, 083042 (2015).

[13]A. B. Poma and L. Delle Site, Phys. Rev. Lett. **104**, 250201 (2010).

[14]A. B. Poma and L. Delle Site, Phys. Chem. Chem. Phys. **13**, 10510 (2011).

[15]R. Potestio and L. Delle Site, J. Chem. Phys. **136**, 054101 (2012).

[16]M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, J. Chem. Phys. **99**, 2796 (1993).

[17]J. Lobaugh and G. A. Voth, J. Chem. Phys. **106**, 2400 (1997).

[18]A. Witt, S. D. Ivanov, M. Shiga, H. Forbert, and D. Marx, J. Chem. Phys. **130**, 194510 (2009).

[19]M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **133**, 124104 (2010).

[20]T. E. Markland and D. E. Manolopoulos, J. Chem. Phys. **129**, 024105 (2008).

[21]M. Ceriotti and D. E. Manolopoulos, Phys. Rev. Lett. **109**, 10064 (2012).

[22]S. Jang, S. Jang, and G. Voth, J. Chem. Phys. **115**, 7832 (2001).

[23]A. Perez and M. E. Tuckerman, J. Chem. Phys. **135**, 064104 (2011).

[24]T. E. Markland and D. E. Manolopoulos, Chem. Phys. Lett. **464**, 256 (2008).

[25]G. S. Fanourgakis, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 094102 (2009).

[26]B. P. Lambeth, C. Junghans, K. Kremer, C. Clementi, and L. Delle Site, J. Chem. Phys. **133**, 221101 (2010).

[27]An important technical point needs to be elucidated which concerns the coarse-grained region. There are two ways to treat the particles in the coarse-grained region: (a) they are point particles which, when entered into the hybrid region, acquire an initial structural and dynamical state, randomly chosen from a distribution taken from the statistics of the PI region; this is the technique used in Ref. 13. However, this approach would not be efficient, for technical reasons, in the current Gromacs code[59] used in this paper. In this case, we use option (b), the beads are retained/carried along into the coarse-grained region and are fixed relative to the path centroid, that is, they are treated as ghost particles which do not interact with beads of other molecules and thus, one reduces the number of calculations for interatomic forces; this corresponds to 75% of the computational cost; thus, the approach is still efficient. However, future work focuses on the optimization with Gromacs of technique (a).

[28]M. Praprotnik, L. Delle Site, and K. Kremer, Phys. Rev. E **73**, 066701 (2006).

[29]M. Praprotnik, L. Delle Site, and K. Kremer, J. Chem. Phys. **126**, 134902 (2007).

[30]M. Praprotnik, S. Matysiak, L. Delle Site, K. Kremer, and C. Clementi, J. Phys.: Condens. Matter **19**, 292201 (2007).

[31]S. Matysiak, C. Clementi, M. Praprotnik, K. Kremer, and L. Delle Site, J. Chem. Phys. **128**, 024503 (2008).

[32]D. Mukherji, N. F. A. van der Vegt, K. Kremer, and L. Delle Site, J. Chem. Theory Comput. **8**, 375 (2012).

[33]M. Praprotnik, K. Kremer, and L. Delle Site, Phys. Rev. E **75**, 017701 (2007).

[34]M. Praprotnik, K. Kremer, and L. Delle Site, J. Phys. A: Math. Gen. **40**, F281 (2007).

[35]J. L. Lebowitz and P. G. Bergmann, Phys. Rev. **99**, 578 (1955).

[36]J. L. Lebowitz and P. G. Bergmann, Ann. Phys. **1**, 1 (1957); see also, J. L. Lebowitz and A. Shimony, Phys. Rev. **128**, 1945 (1962).

[37]M. E. Tuckerman, *Path Integration via Molecular Dynamics*, NIC series Vol. 10 (Forschungszentrum Jülich GmbH, 2002), p. 268.

[38]B. J. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).

[39]D. J. E. Callaway and A. Rahman, Phys. Rev. Lett. **49**, 613 (1982).

[40]M. Parinello and A. Rahman, J. Chem. Phys. **80**, 860 (1984).

[41]I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

[42]R. Potestio, S. Fritsch, P. Espanol, R. Delgado-Buscalioni, K. Kremer, R. Everaers, and D. Donadio, Phys. Rev. Lett. **110**, 108301 (2013).

[43]K. Kreis, D. Donadio, K. Kremer, and R. Potestio, Europhys. Lett. **108**, 30007 (2014).

[44]L. Delle Site, Phys. Rev. E **76**, 047701 (2007).

[45]L. Delle Site, Entropy **16**, 23 (2014).

[46]H. Wang and A. Agarwal, Eur. Phys. J.: Spec. Top. (2015).

[47]R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).

[48]R. Kubo, M. Toda, and N. Hashitsume, in *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer, New York, 1985), Chap. 4.

[49]S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller III, Annu. Rev. Phys. Chem. **64**, 387 (2013).
[50]A. Perez, M. E. Tuckerman, and M. H. Muser, J. Chem. Phys. **130**, 184105 (2009).
[51]S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).
[52]H. Wang, C. Schütte, G. Ciccotti, and L. Delle Site, J. Chem. Theory Comput. **10**, 1376 (2014).
[53]H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622 (2001).
[54]R. M. Lynden Bell and I. R. McDonald, Mol. Phys. **43**, 1429 (1981).
[55]R. W. Impey, P. A. Madden, and I. R. McDonald, Mol. Phys. **46**, 513 (1982).
[56]A. Luzar and D. Chandler, Phys. Rev. Lett. **76**, 928 (1995).
[57]A. Luzar and D. Chandler, Nature **379**, 55 (1996).
[58]F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham III, and G. A. Voth, J. Chem. Phys. **125**, 184507 (2006).
[59]S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, Bioinformatics **29**, 845 (2013).
[60]V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko, J. Chem. Theory Comput. **5**, 3211 (2009).
[61]T. F. Miller III and D. E. Manolopoulos, J. Chem. Phys. **123**, 154504 (2005).