



Linearized symmetrized quantum time correlation functions calculation via phase preaveraging

Michele Monteferrante, Sara Bonella, Giovanni Ciccotti

► To cite this version:

Michele Monteferrante, Sara Bonella, Giovanni Ciccotti. Linearized symmetrized quantum time correlation functions calculation via phase preaveraging. *Molecular Physics*, 2011, pp.1. 10.1080/00268976.2011.619506 . hal-00739205

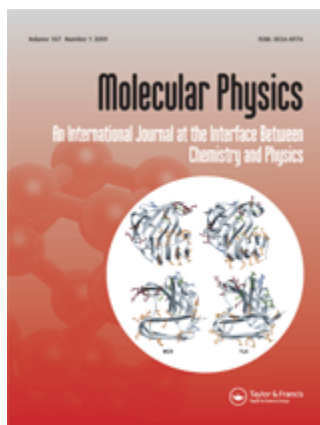
HAL Id: hal-00739205

<https://hal.science/hal-00739205>

Submitted on 6 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Linearized symmetrized quantum time correlation functions calculation via phase preaveraging

Journal:	<i>Molecular Physics</i>
Manuscript ID:	TMPH-2011-0280.R1
Manuscript Type:	Special Issue in honour of Luciano Reatto
Date Submitted by the Author:	23-Aug-2011
Complete List of Authors:	Monteferrante, Michele; Universita' di Roma La Sapienza, Physics Bonella, Sara; Universita' di Roma La Sapienza, Physics Ciccotti, Giovanni; Universita' di Roma La Sapienza, Physics
Keywords:	quantum time correlation functions, semiclassical dynamics, linearized methods, Monte Carlo
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Resub.tex CondIntDublin.bbl	

SCHOLARONE™
Manuscripts

ARTICLE

*Linearized symmetrized quantum time correlation functions calculation via phase preaveraging*M. Monteferrante^{a,*} and S. Bonella^a and G. Ciccotti^{a,b}^a*Dipartimento di Fisica and CNISM Unità 1, Università di Roma “La Sapienza”, Ple A. Moro 2 00185, Rome Italy ;* ^b*School of Physics, Room 302B EMSC-UCD, University College Dublin, Belfield, Dublin 4, Ireland*

(v4.5 released September 2009)

We recently introduced an iterative method to compute quantum time correlation functions [Bonella et al. J. Chem.Phys **133** (16) 164105 (2010)]. There, the thermal part of the correlation function is treated exactly and, similar to the linearization techniques, at zero order of iteration only classical dynamics is required. In this work, we propose a new scheme for the zero order iteration of the method which significantly improves the efficiency of the calculations for high dimensional model systems.

Keywords: approximate quantum time correlation functions; semiclassical dynamics; linearized methods; Monte Carlo algorithm

1. Introduction

Approximate methods to simulate quantum dynamics for high dimensional systems are still limited by numerical instabilities, often due to rapidly oscillating phase factors in the estimators of the observables. An interesting set of schemes that eliminates these phase factors are the so-called linearization methods, usually applied to compute time correlation functions

$$C_{AB}(t, \beta) = \frac{1}{Z} \text{Tr} \left[e^{-\beta \hat{H}} \hat{A} e^{\frac{i}{\hbar} \hat{H} t} \hat{B} e^{-\frac{i}{\hbar} \hat{H} t} \right] \quad (1)$$

where $\hat{A}$ and $\hat{B}$ are the operators of interest, $\hat{H}$ is the Hamiltonian, and $Z = \text{Tr} \left[e^{-\beta \hat{H}} \right]$ the quantum partition function ($\beta = 1/k_B T$). The linearized expression of the time correlation function is (see, for example, [1–5])

$$C_{AB}^l(t, \beta) = \frac{1}{2\pi\hbar} \int dr_0 dp_0 \left[\frac{e^{-\beta \hat{H}}}{Z} \hat{A} \right]_w (r_0, p_0) B_w(r_t, p_t) \quad (2)$$

where (r_t, p_t) are the final points of a classical evolution from the initial conditions (r_0, p_0) , and the subscript w indicates the Wigner transform of the operator, for example $B_w(r_t, p_t) = \int dz e^{\frac{i}{\hbar} p_t z} \langle r_t + z/2 | \hat{B} | r_t - z/2 \rangle$. Linearized methods are useful

*Corresponding author. Email: michele.monteferrante@roma1.infn.it

either for almost classical systems or in situations, not rare in the condensed phase, in which the decay of the correlation functions is fast enough that it can be captured by substituting quantum evolution with classical dynamics, an approximation valid for sufficiently short times also for generic (i.e. non quadratic) potentials. In fact, linearized methods are, at the moment, the only viable approach for systems with more than a few degrees of freedom and have been successfully applied in some interesting calculations, such as, for example, the dynamic structure factor of parahydrogen, [6], and helium above the supercondensation temperature [7], quantum fluctuations in liquid neon [8], and the infrared spectrum of water [9]. Generalizations to the case of non-adiabatic dynamics have also been developed and tested mainly on generalized spin-boson systems [10–12]. In these calculations, eq. (2) is usually interpreted as an average value over a probability density obtained from the absolute value of $\left[\frac{e^{-\beta\hat{H}}}{Z}\hat{A}\right]_w$ and computed via schemes that combine Monte Carlo sampling of the initial conditions (r_0, p_0) and classical molecular dynamics to (r_t, p_t) . The sampling of the initial conditions, however, is non trivial since the Wigner transform of $\hat{A}$ times the density matrix (or even just the Wigner transform of the density matrix) is usually not known. Several approximate schemes exist, based either on the Feynman-Kleinert [13] expression of the density matrix [3], or on developments in Gaussian basis sets [14, 15] of the density matrix or, finally, on procedures employing a generalized Bloch equation to determine the density at low temperatures starting from its, classical, expression for high T [16]. These methods vary in accuracy and numerical efficiency but they all introduce approximations that hinder the usefulness of eq. (2) even for systems for which the use of the linearized dynamics would be justified. **Going beyond strictly linearized methods, there are semiclassical schemes that do not require computing the Wigner transform of the density operator. One interesting example is given by the so called forward-backward dynamics. There a path integral expression (for example in coherent state representation as in [17]) of the density operator is employed for the sampling of initial conditions. While interesting applications of these type of methods exist, also in this case the observable contains phase factors that eventually compromise high dimensional calculations.**

To attempt progress, in [18, 19], we proposed a scheme to compute the symmetrized time correlation function

$$G_{AB}(t, \beta) = \frac{1}{Z} \text{Tr} \left[\hat{A} e^{\frac{i}{\hbar} \hat{H} t_c^*} \hat{B} e^{-\frac{i}{\hbar} \hat{H} t_c} \right] \quad (3)$$

where $t_c = t - i\hbar\beta/2$. **The function above is equivalent, via a relationship among the Fourier transforms, to the standard correlation function and has been considered by many authors as the starting point of approximate schemes of various efficacy [20–28].** Our method offers, in principle, an algorithm that can represent the quantum effects via an iterative scheme. The correct thermal density is sampled at all order of iteration, while quantum dynamical effects are introduced hierarchically with the different orders of iteration. Unfortunately, the numerical cost of the method grows very quickly with the iterations so effective schemes that include higher order terms do not exist and are difficult to conceive. The zero order of the method, however, is interesting. As we shall see in more detail shortly, it requires only classical dynamics to approximate the evolution, but it does not involve the Wigner transform of the density matrix. This opens the possibility to derive a scheme which maintains the advantages of the linearized methods without the problems related to the sampling of the initial conditions. The lowest order approximation of the symmetrized correlation function, see eq. (4), is obtained via a sequence of steps

that, although notationally cumbersome, are not particularly difficult conceptually. The derivation is detailed in [19], here we simply summarize its main points to illustrate the origin of the different sets of variables in eq. (4). To begin with, the trace in eq. (3) is rewritten by introducing representations of the identity in the coordinate basis and isolating the matrix elements of the operators $\hat{A}$ and $\hat{B}$ and of the forward and backward propagators in complex time ($e^{-\frac{i}{\hbar}\hat{H}t_c}$ and $e^{\frac{i}{\hbar}\hat{H}t_c^*}$, respectively). Since $t_c = t - i\hbar\beta/2$, the real time, t , and the “imaginary time” β appear together in the evolution. In our derivation, the real and imaginary propagations are separated (again via resolutions of the identity) and treated differently. Let us begin with the imaginary time evolution. A standard path integral representation is employed first to express it using two sets of ν values of the coordinates (usually referred to as “beads”) of the system, called the forward and backward imaginary paths. The result is then identically rewritten, via a change of variables, in terms of the semi-sum and the difference of the forward and backward imaginary time paths¹. (These new variables are indicated with r^λ and Δr^λ , with $\lambda = 0, \dots, \nu$ in eq. (4).) This form of the imaginary time evolution (which accounts for all the coordinate dependence of the function $\rho(\Gamma)$ in eqns. (4) and (5)) is exact and requires no further manipulation. Let us now move to the forward and backward real time propagators. For these, the so called hybrid path integral representation is employed [3, 29] to represent the evolution in terms of two sets of N values of the momenta and coordinates of the system (the forward and backward real time paths). The hybrid representation is introduced because it simplifies the transition to classical dynamics. Similar to the treatment of imaginary time propagation, the hybrid path integrals are rewritten in terms of semi-sum and difference paths in coordinates and momenta. At this stage, however, a crucial difference is introduced in the treatment of the real time propagation: it is approximated by expanding it to first order in the difference paths. This expansion makes it possible to perform all integrals over the difference variables analytically. The integral over the last difference (real time) bead turns the coordinate representation of operator $\hat{B}$ into the Wigner transform of this operator. More importantly, the integrations over the intermediate (real time) difference beads result in a product of delta functions that effectively reduces the semi-sum paths to a classical trajectory evolving from time 0 to time t . The only integrals relative to the real time path left after use of the delta functions are those on the initial momentum, indicated as p^1 , and the initial coordinate. Due to the concatenation of resolutions of the identity used to isolate real and imaginary time propagators, this coordinate coincides with the end point, r^ν , of the thermal path integral in the semi-sum variables. In addition, the explicit form of the path integral representation of the thermal propagators allows also to perform the integral over the variable Δr^ν analytically (see [19] for details). Thus far, we have focused on the numerator of eq. (3). The denominator can however be manipulated using similar steps to express the approximate symmetrized correlation function as the ratio of two quantities that can, eventually, be computed via Monte Carlo. Keeping all of the above into account, the expression of this ratio is given by

$$G_{AB}^0(t, \beta) = \frac{\int d\Gamma e^{-\frac{i}{\hbar}p^1\Delta r^{(\nu-1)}} \langle r^0 + \frac{\Delta r^0}{2} | \hat{A} | r^0 - \frac{\Delta r^0}{2} \rangle B_W(r_t, p_t) \rho(\Gamma)}{\int d\Gamma e^{-\frac{i}{\hbar}p^1\Delta r^{(\nu-1)}} \rho(\Gamma) \delta(\Delta r^0)} \quad (4)$$

¹if $r_{f/b}^\lambda$ is the λ -th “bead” along the forward/backward path, with $\lambda = 1 \dots \nu$, the change of variables is $r^\lambda = (r_f^\lambda + r_b^\lambda)/2$ and $\Delta r^\lambda = r_f^\lambda - r_b^\lambda$

where $\Gamma = \{p^1, r^0, \dots, r^\nu, \Delta r^0, \dots, \Delta r^{\nu-1}\}$. The function $\rho(\Gamma)$ is defined as

$$\rho(\Gamma) = \frac{1}{Q \sqrt{2\pi\sigma_p^2}} e^{-\frac{(p^1)^2}{2\sigma_p^2}} e^{-\delta_\beta \sum_{\lambda=1}^{\nu} \left[V(r^{(\lambda-1)} + \frac{\Delta r^{(\lambda-1)}}{2}) + V(r^{(\lambda-1)} - \frac{\Delta r^{(\lambda-1)}}{2}) \right]} \times e^{-\frac{\sigma_p^2}{2} \sum_{\lambda=1}^{\nu-1} (\Delta r^\lambda - \Delta r^{(\lambda-1)})^2} e^{-\frac{1}{2\sigma_r^2} \sum_{\lambda=1}^{\nu} (r^\lambda - r^{(\lambda-1)})^2} \quad (5)$$

where Q is its normalization factor, $\delta_\beta = \beta/2(\nu - 1)$, $\sigma_r^2 = \hbar\delta_\beta/2m$, $\sigma_p^2 = m/2\hbar\delta_\beta$. With this definition, $\rho(\Gamma)$ can be interpreted as probability densities and sampled. Similar to eq. (2), (r_t, p_t) in eq. (4) indicates the end point of a purely classical propagation from the initial conditions (r^ν, p^1) . The expression above can then be computed, in analogy with the standard linearized correlation function, via the mixed Monte Carlo - molecular dynamics algorithm that was described in [19]. The form of the density $\rho(\Gamma)$ avoids the need, and the troubles, of the Wigner density, but the price for this is the presence of the phase factor $e^{-\frac{i}{\hbar} p^1 \Delta r^{(\nu-1)}}$ at the numerator and denominator. The tests conducted in [19] indicated that, due to this phase, even the zero order calculation had a very unfavorable scaling with the number of degrees of freedom.

In this work, we describe a new algorithm that improves the convergence of calculations of the zero order approximation to the symmetrized correlation function and scales much better with the number of degrees of freedom. This result opens the possibility to use this algorithm to compute properties of condensed systems with efficiency comparable to that of standard linearized methods, but with higher accuracy. The paper is organized as follows: In section 3, the zero order symmetrized correlation function is rewritten as an average over a more convenient probability density that will eventually be sampled via a new Monte Carlo based algorithm. This probability density is obtained by preaveraging the phase factor in eq. (4). The new probability density, however, does not have an analytic expression but must be calculated numerically. This means that its values are known with statistical uncertainty and standard Monte Carlo cannot be used. We will then show how to compute the average by combining in an original way the Kennedy [30] and the penalty [31] Monte Carlo schemes for sampling "noisy" probability densities. The, non-trivial, new algorithm that results from this combination is described in sections 3 and 4, while the Kennedy and penalty methods are summarized in Appendix I. Finally, in section 5 we test the new method on simple benchmark systems of up to forty degrees of freedom and compare its performance to that of our previous algorithm.

2. Formulation

In the following, we will use one dimensional notation for simplicity (the generalization to the multidimensional case is essentially straightforward). We also assume that the operator $\hat{A}$ is diagonal in the coordinates basis. The case of more general observables, discussed in Appendix II, is more involved notationally but has practically the same complexity. For diagonal $\hat{A}$,

$$\langle r^0 + \frac{\Delta r^0}{2} | \hat{A} | r^0 - \frac{\Delta r^0}{2} \rangle = A(r^0) \delta(\Delta r^0) \quad (6)$$

so the integration over the variable Δr^0 in eq. (4) can be trivially performed. After integration, the correlation function can be written as the ratio of two expectation

values:

$$G_{AB}^0(t, \beta) = \frac{\langle e^{-\frac{i}{\hbar} p^1 \Delta r^{(\nu-1)}} A(r^0) B_W(r_t, p_t) \rangle_{\hat{\rho}(\Gamma')}}{\langle e^{-\frac{i}{\hbar} p^1 \Delta r^{(\nu-1)}} \rangle_{\hat{\rho}(\Gamma')}} \quad (7)$$

where Γ' is the set Γ minus the element Δr^0 , and $\hat{\rho}(\Gamma') = \int d\Delta r^0 \rho(\Gamma) \delta(\Delta r^0) / \hat{Q}$ (with $\hat{Q}$ equal to the normalization over the reduced set of variables). The expression above was used in [19]. To proceed towards the new algorithm, let us define $\mathbf{r} = \{r^0, \dots, r^\nu\}$ and $\Delta \mathbf{r} = \{\Delta r^1, \dots, \Delta r^{\nu-1}\}$. We can isolate convenient terms in the probability density by noting first that it can be identically rewritten as the product of a Gaussian probability density for the momentum and a joint probability density for $\{\mathbf{r}, \Delta \mathbf{r}\}$

$$\hat{\rho}(\Gamma') = \rho_G(p^1) \tilde{\rho}(\mathbf{r}, \Delta \mathbf{r}) \quad (8)$$

with $\rho_G(p^1) = (2\pi\sigma_p^2)^{-1/2} e^{-\frac{(p^1)^2}{2\sigma_p^2}}$ and

$$\begin{aligned} \tilde{\rho}(\mathbf{r}, \Delta \mathbf{r}) = & \frac{1}{\hat{Q}} e^{-2\delta_\beta V(r^0)} e^{-\frac{1}{2\sigma_r^2} \sum_{\lambda=1}^\nu (r^\lambda - r^{(\lambda-1)})^2} \\ & \times e^{-\delta_\beta \sum_{\lambda=2}^\nu \left[V(r^{(\lambda-1)} + \frac{\Delta r^{(\lambda-1)}}{2}) + V(r^{(\lambda-1)} - \frac{\Delta r^{(\lambda-1)}}{2}) \right]} e^{-\frac{\sigma_p^2}{2} \sum_{\lambda=2}^{\nu-1} (\Delta r^\lambda - \Delta r^{(\lambda-1)})^2} e^{-\frac{\sigma_p^2}{2} (\Delta r^1)^2} \end{aligned} \quad (9)$$

We can further decompose the joint probability $\tilde{\rho}(\mathbf{r}, \Delta \mathbf{r})$ as

$$\tilde{\rho}(\mathbf{r}, \Delta \mathbf{r}) = \rho_c(\Delta \mathbf{r} | \mathbf{r}) \rho_m(\mathbf{r}) \quad (10)$$

where ρ_m is the marginal probability density of the variables $\mathbf{r}$ and ρ_c is the conditional probability density of $\Delta \mathbf{r}$ given $\mathbf{r}$. Thus,

$$\begin{aligned} \rho_m(\mathbf{r}) = & \frac{1}{\hat{Q}} e^{-2\delta_\beta V(r^0)} e^{-\frac{1}{2\sigma_r^2} \sum_{\lambda=1}^\nu (r^\lambda - r^{(\lambda-1)})^2} \\ & \times \int d\Delta \mathbf{r} e^{-\delta_\beta \sum_{\lambda=2}^\nu \left[V(r^{(\lambda-1)} + \frac{\Delta r^{(\lambda-1)}}{2}) + V(r^{(\lambda-1)} - \frac{\Delta r^{(\lambda-1)}}{2}) \right]} e^{-\frac{\sigma_p^2}{2} \sum_{\lambda=2}^{\nu-1} (\Delta r^\lambda - \Delta r^{(\lambda-1)})^2} e^{-\frac{\sigma_p^2}{2} (\Delta r^1)^2} \end{aligned} \quad (11)$$

and

$$\begin{aligned} \rho_c(\Delta \mathbf{r} | \mathbf{r}) = & \frac{\tilde{\rho}(\mathbf{r}, \Delta \mathbf{r})}{\rho_m(\mathbf{r})} \\ = & \frac{e^{-\delta_\beta \sum_{\lambda=2}^\nu \left[V(r^{(\lambda-1)} + \frac{\Delta r^{(\lambda-1)}}{2}) + V(r^{(\lambda-1)} - \frac{\Delta r^{(\lambda-1)}}{2}) \right]} e^{-\frac{\sigma_p^2}{2} \sum_{\lambda=2}^{\nu-1} (\Delta r^\lambda - \Delta r^{(\lambda-1)})^2} e^{-\frac{\sigma_p^2}{2} (\Delta r^1)^2}}{\int d\Delta \mathbf{r} e^{-\delta_\beta \sum_{\lambda=2}^\nu \left[V(r^{(\lambda-1)} + \frac{\Delta r^{(\lambda-1)}}{2}) + V(r^{(\lambda-1)} - \frac{\Delta r^{(\lambda-1)}}{2}) \right]} e^{-\frac{\sigma_p^2}{2} \sum_{\lambda=2}^{\nu-1} (\Delta r^\lambda - \Delta r^{(\lambda-1)})^2} e^{-\frac{\sigma_p^2}{2} (\Delta r^1)^2}} \end{aligned} \quad (12)$$

Defining

$$F(p^1, \mathbf{r}) = \int d\Delta \mathbf{r} e^{-\frac{i}{\hbar} p^1 \Delta r^{(\nu-1)}} \rho_c(\Delta \mathbf{r} | \mathbf{r}) \quad (13)$$

and inserting the above factorization of the probability density in the expression of the correlation function we get

$$G_{AB}^0(t, \beta) = \frac{\int dp^1 d\mathbf{r} A(r^0) B_W(r_t, p_t) \rho_m(\mathbf{r}) F(p^1, \mathbf{r}) \rho_G(p^1)}{\int dp^1 d\mathbf{r} \rho_m(\mathbf{r}) F(p^1, \mathbf{r}) \rho_G(p^1)} \quad (14)$$

Let us now consider more in detail the function $F(p^1, \mathbf{r})$: its logarithm is, by definition, the generating function of the cumulants of the random variable $\Delta r^{\nu-1}$ [32]. Indicating the k -th cumulant as $\langle (\Delta r^{\nu-1})^k \rangle_{\rho_c(\Delta \mathbf{r}|\mathbf{r})}^c$, we have in fact

$$\left. \frac{\partial^k \ln F(p^1, \mathbf{r})}{\partial p^{1k}} \right|_{p^1=0} = \left(-\frac{i}{\hbar} \right)^k \langle (\Delta r^{\nu-1})^k \rangle_{\rho_c(\Delta \mathbf{r}|\mathbf{r})}^c \quad (15)$$

from which it also follows

$$\ln F(p^1, \mathbf{r}) = \sum_{n=1}^{\infty} \frac{(-ip^1/\hbar)^n}{n!} \langle (\Delta r^{\nu-1})^n \rangle_{\rho_c(\Delta \mathbf{r}|\mathbf{r})}^c \quad (16)$$

The function $F(p^1, \mathbf{r})$ can then be formally expressed as

$$F(p^1, \mathbf{r}) = \exp \left[\sum_{n=1}^{\infty} \frac{(-ip^1/\hbar)^n}{n!} \langle (\Delta r^{\nu-1})^n \rangle_{\rho_c(\Delta \mathbf{r}|\mathbf{r})}^c \right] \quad (17)$$

The conditional probability $\rho_c(\Delta \mathbf{r}|\mathbf{r})$, eq. (12), is an even function of $\Delta \mathbf{r}$. This implies that the odd terms in the cumulant expansion are equal to zero. $F(p^1, \mathbf{r})$ is then a real function and can be written as

$$F(p^1, \mathbf{r}) = e^{-E(p^1, \mathbf{r})} \quad (18)$$

where we defined

$$E(p^1, \mathbf{r}) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} (p^1/\hbar)^{2n}}{(2n)!} \langle (\Delta r^{\nu-1})^{2n} \rangle_{\rho_c(\Delta \mathbf{r}|\mathbf{r})}^c \quad (19)$$

Thus, if the cumulant expansion above converges, $F(p^1, \mathbf{r})$ is a positive definite function that can be used, together with $\rho_G(p^1)$ and $\rho_m(\mathbf{r})$ to define the probability density

$$\mathcal{P}(p^1, \mathbf{r}) = \frac{\rho_G(p^1) e^{-E(p^1, \mathbf{r})} \rho_m(\mathbf{r})}{\int dp^1 d\mathbf{r} \rho_G(p^1) e^{-E(p^1, \mathbf{r})} \rho_m(\mathbf{r})} \quad (20)$$

With this definition the zero order symmetrized correlation function can be written as

$$G_{AB}^0(t, \beta) = \langle A(r^0) B_W(r_t, p_t) \rangle_{\mathcal{P}} \quad (21)$$

The average value defined above has the nice characteristic that no phase factors are present in the observable. The convergence of a Monte Carlo scheme based on sampling $\mathcal{P}$ should then be comparable to that of standard calculations. On the other hand, standard Monte Carlo cannot be applied to sample this probability

which can only be computed approximately. In the next section we discuss how to overcome this problem.

3. Method

Let us start recalling that Monte Carlo samples a preassigned probability density $P(s)$ (s is a state of the system) by identifying a Markov process that performs a random walk through the states of the system based on a transition probability $\Pi(s \rightarrow s')$. If the process is ergodic and detailed balance is satisfied, i.e. if $P(s)\Pi(s \rightarrow s') = P(s')\Pi(s' \rightarrow s)$, the random walk will asymptotically sample $P(s)$. The most common way to ensure detailed balance, the Metropolis method, writes the transition probability as the product of the (preassigned and usually symmetric) probability to generate state s' given that we are in s , $T(s \rightarrow s')$, times the acceptance probability $A(s \rightarrow s')$ and defines

$$A(s \rightarrow s') = \min \left[1, \frac{P(s')T(s' \rightarrow s)}{P(s)T(s \rightarrow s')} \right] \tag{22}$$

For future reference note that if, as it is often the case, the probability is of the form $P(s) \propto e^{-L(s)}$, the equation above can be rewritten in the form

$$A(s \rightarrow s') = \min \left[1, \frac{e^{-\Delta(s,s')}T(s' \rightarrow s)}{T(s \rightarrow s')} \right] \tag{23}$$

with $\Delta(s, s') = L(s') - L(s)$. The Metropolis scheme is well established when the asymptotic distribution is known analytically. However, if $P(s)$ has to be estimated via some numerical method, the noise associated with the estimate can bias the acceptance probability and compromise sampling. This is precisely the case for $\mathcal{P}(p^1, \mathbf{r})$ since the functions $E(p^1, \mathbf{r})$ and $\rho_m(\mathbf{r})$ must be calculated via eqns. (19) and (11), respectively. An effective strategy suggested in the literature to address this problem is to modify the acceptance probability so as to account for the effect of the noise. Two main algorithms, known as the Kennedy [30] and penalty [31] methods, exist for that purpose. Due to the specific form of our probability, none of these methods can be directly applied, but they can be combined to provide an original Monte Carlo algorithm suitable to our purposes. Before doing that, we quickly recall these schemes and introduce some definitions. A brief presentation of [30] and [31] can be found in Appendix I.

The Kennedy method (see Appendix I) was devised to sample a probability density the form $P(s) \propto e^{-L(s)}f(s)$, where $L(s)$ is an analytically known function, and $f(s) \geq 0$ is a function with noise. Assuming that an unbiased estimator, $U(s \rightarrow s')$, of the ratio $f(s')/f(s)$ exists, Kennedy et al. showed that a suitable Monte Carlo can be constructed by choosing the probability to generate the move, $T_K(s \rightarrow s')$, and the acceptance probability, $A_K(s \rightarrow s')$, as

$$T_K(s \rightarrow s') \propto e^{-L(s')} \tag{24}$$

and

$$A_K(s \rightarrow s') = \begin{cases} c_+ + c_- U(s \rightarrow s') & \text{if } "s > s'" \\ c_- + c_+ U(s' \rightarrow s) & \text{if } "s \leq s'" \end{cases} \tag{25}$$

The parameters c_+, c_- in the definition above must ensure that $A_K(s \rightarrow s')$ only has values in the interval $[0, 1]$. The conditions “ $s > s'$ ”, “ $s \leq s'$ ” imply the choice of an ordering criterion of the configurations s, s' . The best choice of this criterion depends on the problem, but the one usually adopted is “ $s > s'$ ” equivalent to $e^{-L(s)} > e^{-L(s')}$. An auxiliary, standard, Monte Carlo is used to generate states according to eq. (24).

The penalty method (see Appendix I), on the other hand, was developed to sample probability densities of the form $P(s) \propto e^{-L(s)}$ when the function $\Delta(s, s')$ in eq. (23) is too expensive to be computed exactly at each Monte Carlo move. Indicating with $\delta_i(s, s')$ a specific value of the estimate of $\Delta(s, s')$, and with

$$D(s, s') = \frac{1}{N_s} \sum_{i=1}^{N_s} \delta_i(s, s')$$

$$\chi^2(s, s') = \frac{1}{N_s(N_s - 1)} \sum_{i=1}^{N_s} (D - \delta_i(s, s'))^2 \quad (26)$$

the estimates of the mean and of the variance associated to N_s evaluations, in [31] it was shown that Monte Carlo sampling of $P(s)$ can be performed using the acceptance probability

$$a_P(D(s, s'), \chi^2) = \min \left[1, e^{-D(s, s') - u_{\chi^2}} \right] \quad (27)$$

where the function u_{χ^2} has the following expansion in χ^2 for $\chi^2/N_s < 1/4$:

$$u_{\chi^2} = \frac{\chi^2}{2} + \frac{\chi^4}{4(N_s + 1)} + \frac{\chi^6}{3(N_s + 1)(N_s + 3)} + \dots \quad (28)$$

Let us now go back to the calculation of the average in eq. (21). To pave the way for our algorithm, we rewrite $\mathcal{P}(p^1, \mathbf{r})$ in a convenient form. Define

$$e^{-V_r(\mathbf{r})} \doteq e^{-\frac{1}{2\sigma_r^2} \sum_{\lambda=1}^{\nu} (r^\lambda - r^{(\lambda-1)})^2},$$

$$e^{-V_\Delta(\Delta\mathbf{r})} \doteq e^{-\frac{\sigma_p^2}{2} \sum_{\lambda=2}^{\nu-1} (\Delta r^\lambda - \Delta r^{(\lambda-1)})^2} e^{-\frac{\sigma_p^2}{2} (\Delta r^1)^2}, \quad (29)$$

$$e^{-\delta_\beta \bar{V}(\mathbf{r}, \Delta\mathbf{r})} \doteq e^{-\delta_\beta \sum_{\lambda=1}^{\nu} \left[V(r^{(\lambda-1)} + \frac{\Delta r}{2} r^{(\lambda-1)}) + V(r^{(\lambda-1)} - \frac{\Delta r}{2} r^{(\lambda-1)}) \right]} e^{-2\delta_\beta V(r^0)}$$

Substituting these definitions in eq. (11), $\rho_m(\mathbf{r})$ can be written as

$$\rho_m(\mathbf{r}) = \int d\Delta\mathbf{r} \tilde{\rho}(\mathbf{r}, \Delta\mathbf{r})$$

$$= \frac{1}{\hat{Q}} e^{-V_r(\mathbf{r})} \int d\Delta\mathbf{r} e^{-\delta_\beta \bar{V}(\mathbf{r}, \Delta\mathbf{r})} e^{-V_\Delta(\Delta\mathbf{r})} \quad (30)$$

$$\doteq \frac{1}{\hat{Q}} e^{-V_r(\mathbf{r})} \rho'_m(\mathbf{r})$$

where the last line defines $\rho'_m(\mathbf{r})$. With this notation,

$$\mathcal{P}(p^1, \mathbf{r}) = \frac{1}{\hat{Q}} \rho_G(p^1) e^{-E(p^1, \mathbf{r})} e^{-V_r(\mathbf{r})} \rho'_m(\mathbf{r}) \quad (31)$$

Note that the variables p^1 and $\mathbf{r}$ are not independent so they must be treated together when sampling $\mathcal{P}$. In the next section we describe a Monte Carlo suitable to this purpose.

4. Algorithm

The Monte Carlo to sample $\mathcal{P}(p^1, \mathbf{r})$ is constructed as follows. Begin by choosing, with probability 1/2, if the state of the system is changed by attempting to sample a new momentum or a new (set of) coordinate(s).

i) A move for the momentum p^1 has been selected. In this case, detailed balance takes the form

$$\begin{aligned} \mathcal{P}(p^1, \mathbf{r}) T^p(p^1 \rightarrow p^{1'}) A^p((p^1, \mathbf{r}) \rightarrow (p^{1'}, \mathbf{r})) \\ = \mathcal{P}(p^{1'}, \mathbf{r}) T^p(p^{1'} \rightarrow p^1) A^p((p^{1'}, \mathbf{r}) \rightarrow (p^1, \mathbf{r})) \end{aligned} \quad (32)$$

Define the probability to generate a move as

$$T^p(p^1 \rightarrow p^{1'}) = \rho_G(p^{1'}) \quad (33)$$

Using the explicit expression of $\mathcal{P}$, detailed balance can be simplified as

$$e^{-E(p^1, \mathbf{r})} A^p((p^1, \mathbf{r}) \rightarrow (p^{1'}, \mathbf{r})) = e^{-E(p^{1'}, \mathbf{r})} A^p((p^{1'}, \mathbf{r}) \rightarrow (p^1, \mathbf{r})) \quad (34)$$

where $E(p^1, \mathbf{r})$ (and $E(p^{1'}, \mathbf{r})$) is evaluated via the estimator defined below. With the identification $L = E(p^1, \mathbf{r})$, the relation above has the same structure of detailed balance in the penalty method, which requires an appropriate estimate, D_p , of the difference $E(p^1, \mathbf{r}) - E(p^{1'}, \mathbf{r})$ and of χ^2 (see equation (26)), to compute the acceptance probability

$$a_p^p(p^1 \rightarrow p^{1'}) = \min \left[1, e^{-D_p - u_{\chi^2}} \right] \quad (35)$$

and accept and reject moves. To construct D_p , and then χ^2 , note that the different terms in the definition of $E(p^1, \mathbf{r})$, eq. (19), are averages over the conditional probability density $\rho_c(\Delta \mathbf{r} | \mathbf{r})$. $E(p^1, \mathbf{r})$ can be estimated, sampling via an auxiliary Monte Carlo $\rho_c(\Delta \mathbf{r} | \mathbf{r})$, as

$$\tilde{E}(p^1, \mathbf{r}) = \sum_{n=1}^{n_{\max}} \frac{(-1)^{n-1} (p^1 / \hbar)^{2n}}{(2n)!} G_{2n} \quad (36)$$

where G_{2n} is an unbiased estimator of the cumulant of order $2n$. For example, remembering that the average value of $\Delta r^{\nu-1}$ is zero, $G_2 = \sum_{i=1}^N (\Delta r_i^{\nu-1})^2 / N$. Computing also $\tilde{E}(p^{1'}, \mathbf{r})$ we get $\delta = \tilde{E}(p^{1'}, \mathbf{r}) - \tilde{E}(p^1, \mathbf{r})$. Given a set of N_s estimates of this difference, both D_p and χ^2 can be calculated.

To sample $\rho_c(\Delta \mathbf{r} | \mathbf{r})$ we use

$$T(\Delta \mathbf{r} \rightarrow \Delta \mathbf{r}') \propto e^{-V_\Delta(\Delta \mathbf{r}')} \quad (37)$$

to generate the new configurations given the old ones ¹ and accept or reject moves

¹ $T(\Delta \mathbf{r} \rightarrow \Delta \mathbf{r}')$ is a product of Gaussians that can be sampled directly using relatively standard methods

based on

$$A(\Delta \mathbf{r} \rightarrow \Delta \mathbf{r}') = \min \left[1, e^{-\delta_\beta (\bar{V}(\mathbf{r}, \Delta \mathbf{r}') - \bar{V}(\mathbf{r}, \Delta \mathbf{r}))} \right] \quad (38)$$

ii) A move for the (set of) coordinate(s) $\mathbf{r}$ has been selected. Detailed balance reads

$$\begin{aligned} \mathcal{P}(p^1, \mathbf{r}) T^r(\mathbf{r} \rightarrow \mathbf{r}') A^r((p^1, \mathbf{r}) \rightarrow (p^1, \mathbf{r}')) \\ = \mathcal{P}(p^1, \mathbf{r}') T^r(\mathbf{r}' \rightarrow \mathbf{r}) A^r((p^1, \mathbf{r}') \rightarrow (p^1, \mathbf{r})) \end{aligned} \quad (39)$$

which, using the explicit definition of $\mathcal{P}$, becomes

$$\begin{aligned} e^{-E(p^1, \mathbf{r})} e^{-V_r(\mathbf{r}')} \rho'_m(\mathbf{r}) T^r(\mathbf{r} \rightarrow \mathbf{r}') A^r((p^1, \mathbf{r}) \rightarrow (p^1, \mathbf{r}')) \\ = e^{-E(p^1, \mathbf{r}')} e^{-V_r(\mathbf{r})} \rho'_m(\mathbf{r}') T^r(\mathbf{r}' \rightarrow \mathbf{r}) A^r((p^1, \mathbf{r}') \rightarrow (p^1, \mathbf{r})) \end{aligned} \quad (40)$$

If $E(p^1, \mathbf{r})$ were known exactly, this relation would correspond to detailed balance in the Kennedy method via the identification $L = E + V_r$ and $f = \rho'_m$ (see Appendix I). We could then satisfy eq. (40) generating moves according to the probability

$$T^r(\mathbf{r} \rightarrow \mathbf{r}') \propto e^{-E(p^1, \mathbf{r}')} e^{-V_r(\mathbf{r}')} \quad (41)$$

and using the acceptance probability of eq. (25), provided that an unbiased estimator for the ratio $\rho'_m(\mathbf{r}')/\rho'_m(\mathbf{r})$ can be defined. In Kennedy, sampling of the probability density in eq. (41) is realized via an auxiliary, standard, Monte Carlo. When this Markov chain reaches the asymptotic distribution, the configurations generated can be used in the Kennedy acceptance test. In our case this procedure cannot be applied directly since the function L is given by the sum of an analytic term, $V_r(\mathbf{r})$, and of the noisy function $E(p^1, \mathbf{r})$. We can however tackle the problem of generating trial moves distributed as T^r by substituting the penalty procedure to standard Monte Carlo in the auxiliary calculation. To that end, observe that the detailed balance for sampling $e^{-[E+V_r]}$

$$\begin{aligned} e^{-E(p^1, \mathbf{r})} e^{-V_r(\mathbf{r})} t(\mathbf{r} \rightarrow \mathbf{r}') A_P^r((p^1, \mathbf{r}) \rightarrow (p^1, \mathbf{r}')) \\ = e^{-E(p^1, \mathbf{r}')} e^{-V_r(\mathbf{r}')} t(\mathbf{r}' \rightarrow \mathbf{r}) A_P^r((p^1, \mathbf{r}') \rightarrow (p^1, \mathbf{r})) \end{aligned} \quad (42)$$

is satisfied choosing

$$t(\mathbf{r} \rightarrow \mathbf{r}') \propto e^{-V_r(\mathbf{r}')} \quad (43)$$

(this function is a product of Gaussians that can be sampled directly) and using the penalty acceptance

$$a_P^r(\mathbf{r} \rightarrow \mathbf{r}') = \min \left[1, e^{-D_r - \tilde{u}_{\chi^2}} \right] \quad (44)$$

to accept or reject moves. In (44), D_r is the estimator of the difference $E(p^1, \mathbf{r}') - E(p^1, \mathbf{r})$ and $\tilde{u}_{\chi^2}$ is defined as in eq. (28). They are obtained, via eqns. (26), using N_s realizations of the stochastic variable $\delta_r = \tilde{E}(p^1, \mathbf{r}') - \tilde{E}(p^1, \mathbf{r})$. Note that each estimate of δ_r requires two Monte Carlo calculations: $\rho_c(\Delta \mathbf{r} | \mathbf{r})$ to compute $\tilde{E}(p^1, \mathbf{r})$

such as staging [33] or the Levy flight algorithm [34].

and $\rho_c(\Delta\mathbf{r}|\mathbf{r}')$ to compute $\tilde{E}(p^1, \mathbf{r}')$. To proceed with the Kennedy acceptance test, we must provide an unbiased estimator for the ratio $\rho'_m(\mathbf{r}')/\rho'_m(\mathbf{r})$. Since

$$\begin{aligned} \frac{\rho'_m(\mathbf{r}')}{\rho'_m(\mathbf{r})} &= \frac{\int d\Delta\mathbf{r} e^{-V_\Delta(\Delta\mathbf{r})} e^{-\delta_\beta \bar{V}(\mathbf{r}', \Delta\mathbf{r})}}{\int d\Delta\mathbf{r} e^{-V_\Delta(\Delta\mathbf{r})} e^{-\delta_\beta \bar{V}(\mathbf{r}, \Delta\mathbf{r})}} \\ &= \frac{\int d\Delta\mathbf{r} e^{-V_\Delta(\Delta\mathbf{r})} e^{-\delta_\beta \bar{V}(\mathbf{r}, \Delta\mathbf{r})} e^{-\delta_\beta [\bar{V}(\mathbf{r}', \Delta\mathbf{r}) - \bar{V}(\mathbf{r}, \Delta\mathbf{r})]}}{\int d\Delta\mathbf{r} e^{-V_\Delta(\Delta\mathbf{r})} e^{-\delta_\beta \bar{V}(\mathbf{r}, \Delta\mathbf{r})}} \\ &= \langle e^{-\delta_\beta [\bar{V}(\mathbf{r}', \Delta\mathbf{r}) - \bar{V}(\mathbf{r}, \Delta\mathbf{r})]} \rangle_{\rho_c(\Delta\mathbf{r}|\mathbf{r})} \end{aligned} \quad (45)$$

an unbiased estimator is

$$U(\mathbf{r} \rightarrow \mathbf{r}') = \frac{1}{N} \sum_{i=1}^N e^{-\delta_\beta [\bar{V}(\mathbf{r}', \Delta\mathbf{r}_i) - \bar{V}(\mathbf{r}, \Delta\mathbf{r}_i)]} \quad (46)$$

In the equation above, $\{\Delta\mathbf{r}_i\}$ is a sample of size N , with N large enough to guarantee good convergence, distributed as $\rho_c(\Delta\mathbf{r}|\mathbf{r})$ and calculated using the Monte Carlo that employs eqns. (37) and (38)¹. Given $U(\mathbf{r} \rightarrow \mathbf{r}')$, moves are accepted or rejected according to Kennedy's acceptance probability, i.e.

$$A_K^r(\mathbf{r} \rightarrow \mathbf{r}') = \begin{cases} c_+ + c_- U(\mathbf{r} \rightarrow \mathbf{r}') & \text{if " } \mathbf{r} > \mathbf{r}' \text{ " } \\ c_- + c_+ U(\mathbf{r}' \rightarrow \mathbf{r}) & \text{if " } \mathbf{r} \leq \mathbf{r}' \text{ " } \end{cases} \quad (47)$$

Implementing eq. (47), we adopted the prescription, suggested in [30], $c_+ = 0$, $c_- = c < 1$. This is a convenient choice since it allows to compute the estimator U only when " $\mathbf{r} > \mathbf{r}'$ ". The choice of the ordering criterion also mimics [30] and we set: " $\mathbf{r} > \mathbf{r}'$ equivalent to $e^{-\delta_\beta \bar{V}(\Delta\mathbf{r}=0, \mathbf{r})} > e^{-\delta_\beta \bar{V}(\Delta\mathbf{r}=0, \mathbf{r}')}$ ".

Note that, due to the auxiliary Monte Carlo calculations necessary to compute $\rho'_m(\mathbf{r}')/\rho'_m(\mathbf{r})$ and/or $D_{r,p}$, the total number, N_T , of moves in this algorithm is given by the number of steps in the main Monte Carlo, indicated as N_{MC} in the following, times the number of steps of the auxiliary Monte Carlo, indicated with N .

5. Results

We test our algorithm on two typical benchmark models [35] for which exact results can be obtained via basis set expansion method (see [19] for details): the multidimensional harmonic oscillator and the quartic potential in the Hamiltonians eq.(49) and eq.(50). These systems were also employed as benchmarks for the first algorithm we proposed for the symmetrized correlation function. For both systems, we compute the position-position correlation function $G_{rr}(t, \beta)$

$$G_{rr}(t, \beta) = \frac{1}{Z} \text{Tr} \left\{ \sum_{j=1}^d \hat{r}_j e^{i\hat{H}t_c} \hat{r}_j^* e^{-i\hat{H}t_c} \right\} \quad (48)$$

¹Given the exponential form of the observable, the convergence of this calculation might be difficult. However, since the configurations $\mathbf{r}$ and $\mathbf{r}'$ are close (in the spirit of a standard Metropolis move), this should not be a problem.

for a particle with unitary mass (arbitrary units, with $\hbar = k_B = 1$, will be used in this section). Let us begin with a system of d noninteracting harmonic oscillators with Hamiltonian

$$\hat{H}_h = \sum_{i=1}^d \left[\frac{\hat{p}_i^2}{2} + \frac{\hat{r}_i^2}{2} \right] \quad (49)$$

at a temperature $T = 1.0$. As it is well known, for this system eq. (4), and in general linearized methods, are exact so any discrepancy between the quantum result and the results obtained with approximate methods is due to statistical noise. We performed a set of calculations for increasing values of d to compare the convergence properties of the new algorithm with those of the calculation of our previous work. The new algorithm requires to fix three parameters: ν , the number of beads in the path integral expression of the density matrix element (see eq. (5)), the parameter c in the Kennedy acceptance probability (see eq. (47)), and n_{max} , the number of terms in $\tilde{E}(p^1, \mathbf{r})$ (see eq. (36)). For harmonic systems, however, $\mathbf{r}$ and $\Delta \mathbf{r}$ are in fact independent random variables and the probability density for the $\Delta \mathbf{r}$ is a product of Gaussians. In this case, the cumulant expansion terminates, analytically, at second order. A possible convergence tests on the function $\tilde{E}$ for non harmonic systems is mentioned in our next test calculation. Moving on, ν can be chosen (independently of the other parameters) via a preliminary path integral calculation of equilibrium, time independent, properties. For all systems studied in this paper, we performed these calculations using a standard path integral code and we fixed ν by computing the average of the position and of the square of the position. In this first harmonic test, these quantities were converged with $\nu = 12$ beads. The choice of the parameter c is more delicate. Its value is determined by finding (numerically) a good compromise among two requirements: (1) c must be small enough to guarantee that the r.h.s. of eq.(47) (being a probability) is less than one; (2) c must be large enough to ensure a reasonable acceptance of the trial moves in the Monte Carlo. Satisfying these requirements simultaneously may be difficult. The most problematic case is when $U(\mathbf{r} \rightarrow \mathbf{r}')$ is not bound everywhere, so that it may not be possible to find a value of c that will satisfy condition (1) for all configurations. While this is a known problem of the method, Kennedy showed empirically in [30] that violations of this condition of order of one configuration in a thousand or less do not affect the algorithm. For the harmonic systems considered here, we found that $c = 0.9$ is a good choice. With this value we observed no violations of condition (1) (referred to as the limiting condition in the following), and accepted about 40% of the Kennedy moves for all values of d . In Table 1 we compare the performance of the old and new algorithm for $G_{rr}^0(t = 3.14, \beta = 1)$, results for different times show the same behaviour. In the Table, G_{rr}^{ex} indicates the exact results and $G_{rr}^{[1]}$ those obtained with the old algorithm, while $N_T = N \times n_{MC}$ is the total number of Monte Carlo steps necessary to converge to the exact result with an accuracy of about 5% with the new algorithm. $N_T^{[1]}$ is the number of Monte Carlo steps necessary to converge, with the same accuracy, with the algorithm used in [19]. The data shows that the old algorithm is more efficient than the new one for $d \leq 5$. This is due to the computational cost of the auxiliary Monte Carlo calculations that we now have to perform which, for low dimensional systems, is not compensated by the absence of the phase in the observable. However, already for $d = 10$, the new algorithm is about 3 times faster than the old one, and it gains a factor of about 10 for the higher dimensional cases. The advantage becomes more relevant if a higher precision is required. For example, if the error is reduced to 1% of the exact result, calculations with the old algorithm require about 40 times

more Monte Carlo moves for 10 degrees of freedom.

In [19], we showed that the bad scaling of the original algorithm for the zero order correlation function could be mitigated considerably via a stationary phase approximation. In fact, for harmonic systems, we demonstrated that this approximation reduces the scaling to essentially linear in d and does not affect the dynamics of the system appreciably. Unfortunately, for general potentials, after stationary phase, the time evolved correlation function differs from the one corresponding to the fully linearized G_{AB}^0 . To see that, let us consider the system

$$\hat{H}_q = \frac{\hat{p}^2}{2} + \frac{\hat{r}^4}{4} \quad (50)$$

with $T = 1$ (this unidimensional case is already enough to show the problem that we wish to discuss). In the left panel of figure 1 we compare $G_{rr}^{ex}(t, 1)$ computed via a converged basis set expansion (our exact benchmark, represented as the red curve) with $G_{AB}^0(t, 1)$ calculated using the original algorithm (green crosses) or the algorithm that includes the stationary phase (blue triangles) approximation. The blue curve deviates almost immediately both from the exact result and from the zero order approximation computed via the old algorithm, which is able to reproduce the red curve up to $t \approx 2$. The comparison of the full zero order approximation (green crosses) and of the result obtained with our new algorithm (magenta dots), on the other hand, is shown in the right panel of the figure, where we also report for reference the exact result (red curve). The green and magenta symbols are essentially superimposed at all times showing that, contrary to the stationary phase, the new algorithm does not disturb the time evolution. For this system we employed the same values of ν and c as in the harmonic case (Kennedy acceptance probability about 50% and about one violation of the limiting condition every 10^6 moves) and we truncated the cumulant expansion in $E(p^1, \mathbf{r})$ at second order. For non harmonic systems this truncation is not exact, however we tested it by computing the correlation function at $t = 0$ first with $n_{max} = 1$ and then with $n_{max} = 2$ in eq. (36) and verifying that the results of these calculations were equal within error bars.

We used the quartic system also to verify that the favorable scaling of the algorithm with dimensions is maintained for non harmonic models by looking at the system $\hat{H}_q^d = \sum_{i=1}^d [\hat{p}_i^2/2 + \hat{r}_i^4/4]$. In Table 2 we report results for $d = 1, \dots, 40$. As before, G_{rr}^{ex} indicates the exact result and N_T is the total number of steps necessary to converge to within 5% of the exact result (in this case we do not report the results with the old algorithm since the bad scaling observed in the harmonic calculation does not change). The Table shows that the performance of the algorithm is very similar to the harmonic case for $d < 10$. The, slight, increase in the cost for higher dimensions, which is due mainly to the cost of computing accurately enough the cumulants for this non-harmonic system, is still acceptable and well below the estimated behavior of the old algorithm.

As a final check of the method, in Table 3 we report results for the quartic system at the lower temperature $T = 0.2$. For these calculations, convergence requires to increase the number of beads in the thermal path integral to $\nu = 24$, and we found that $c = 0.6$ is a convenient choice in the Kennedy acceptance probability (acceptance probability of about 50% and about one violation of the limiting condition every 10^6 moves). Also in this case the cumulant expansion in $\tilde{E}(p^1, \mathbf{r})$ was converged for $n_{max} = 1$. While the cost of the calculations increases, as do all path integral based calculations, with decreasing temperature, it is still much cheaper than the previous algorithm. For comparison, not shown in the Table, for

$d = 10$ the old algorithm was not converged after 10^9 steps.

6. Conclusions

In this paper, we presented a new algorithm for calculating the zero order expression of the symmetrized quantum time correlation functions introduced in [19]. By performing test calculations for model systems with up to 40 degrees of freedom, we demonstrated that the new algorithm is considerably more efficient than the brute force Monte Carlo we used in that work. As discussed in [19], the approximate symmetrized correlation function is as accurate as the so-called Wigner or linearized form of quantum correlation functions. Both our and these methods sample the exact thermal density and approximate the quantum time propagation via purely classical dynamics. Linearized methods, however, require sampling the Wigner transform of the density matrix (times an operator), a task that cannot be performed exactly for general high dimensional systems. While approximate methods to do this sampling exists, they increase the cost of linearized calculations and often introduce further approximations in the correlation function. Our scheme, on the other hand, employs a path integral representation of the thermal density that can be sampled exactly. Memory of the Wigner transform is carried in the form of a phase factor that had, so far, hindered high dimensional applications of the method. The new algorithm domesticates the phase oscillations via preaveraging and achieves reasonable scaling with dimensionality without loss of accuracy. The results obtained so far are quite encouraging and open the way to applications to realistic condensed phase problems. Moreover, work is in progress to extend the method to compute higher order iterations of the method presented in [19].

7. Acknowledgements

This paper is dedicated to Luciano Reatto, on the occasion of his retirement, as a testimony of his tireless efforts to convince us classical statistical mechanics simulators to study quantum problems. The authors are grateful to D. Ceperley and E. Liberatore for useful discussions on the penalty method. Financial support from SFI Grant No. 08-IN.1-I1869 and from the Istituto Italiano di Tecnologia under the SEED project grant No. 259 SIMBEDD - Advanced Computational Methods for Biophysics, Drug Design and Energy Research is acknowledged.

d	G_{rr}^{ex}	$G_{rr}^{[1]}$	G_{rr}^0	$N_{MC} \times 500$	N	$N_T \times 10^5$	$N_T^{[1]} \times 10^5$
1	-0.96	-0.98 ± 0.05	-0.99 ± 0.05	18	140	12.6	0.25
5	-4.8	-4.6 ± 0.2	-4.9 ± 0.2	22	200	22	15
10	-9.6	-9.7 ± 0.5	-9.5 ± 0.5	23	200	23	75
20	-19.2	-18.8 ± 0.7	-19.0 ± 0.7	52	300	78	1000
40	-38.4	-37 ± 2	-40 ± 2	101	600	303	2500

Table 1. Position autocorrelation function for the harmonic oscillator with unitary mass for different values of the system dimension d . G_{rr}^{ex} are the exact results and $G_{rr}^{[1]}$ and G_{rr}^0 the results obtained with the algorithm introduced in [19] and in this work, respectively. All correlation functions are computed at $t = 3.14$ and $\beta = 1$. N_T is the total number of Monte Carlo steps which, for the new algorithm, is given by the product of the $(\mathbf{r}, p^1)$ moves, N_{MC} , times the number of steps, N , needed to estimate $E(p^1, \mathbf{r})$ and $\rho'_m(\mathbf{r})$. In the last column on the right we report the total number of steps used in [19]. In all calculations, the total number of steps was fixed to obtain a result within about 5% of the exact result.

d	G_{rr}^{ex}	G_{rr}^0	$N_{MC} \times 500$	N	$N_T \times 10^5$
1	0.58	0.58 ± 0.03	18	100	9
5	2.9	2.8 ± 0.2	20	200	20
10	5.8	5.7 ± 0.3	30	200	30
20	11.5	11.5 ± 0.6	46	600	138
40	23	24 ± 1	100	1200	600

Table 2. Quartic potential with unitary mass and inverse temperature for different values of the system dimension d . Symbols are as in Table. 1 but calculations were performed at $t = 0$ (note that the phase factor affects time zero and subsequent times in the same way so even $t = 0$ quantities can be used to test the new algorithm). The total number of steps in the table was fixed to obtain an error of about 5% with respect to the exact result.

d	G_{rr}^{ex}	G_{rr}^0	$N_{MC} \times 500$	N	$N_T \times 10^5$
1	0.060	0.061 ± 0.007	200	200	200
5	0.30	0.27 ± 0.04	200	300	300
10	0.60	0.58 ± 0.03	1800	600	5400
20	1.2	1.15 ± 0.06	2000	800	8000
40	2.4	2.3 ± 0.1	2000	2000	20000

Table 3. Results for the quartic potential at the lower temperature $T = 0.2$. As before, the number of steps was fixed to obtain an error of about 5% with respect to the exact result.

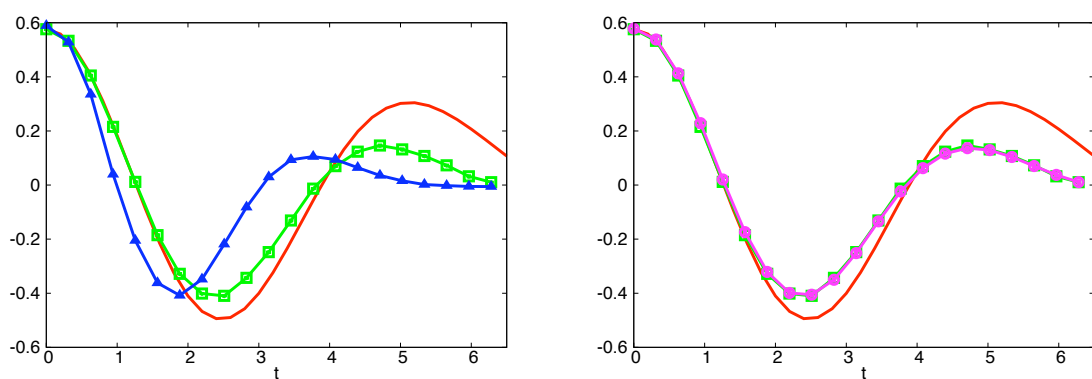


Figure 1. Position auto correlation function for a one dimensional quartic potential with unitary mass and inverse temperature. In both panels, we show in red (curve with no symbols) the exact result and in green (curve with the empty squares) the one obtained with the method used in [19]. The blue curve with triangles in the left panel is the result obtained via the stationary phase approximation. In the right panel, we have added to red and green curves the results obtained with the algorithm (magenta curve with filled circles) introduced in the present work. The green and magenta curves are superimposed.

Appendix I

Kennedy

The definitions in eq. (24) and (25) are justified showing that, asymptotically, the detailed balance relation

$$P(s)T_K(s \rightarrow s')A_K(s \rightarrow s') - P(s')T_K(s' \rightarrow s)A_K(s' \rightarrow s) = 0 \quad (51)$$

is satisfied. In fact, using eq. (24) and $P(s) \propto e^{-L(s)}f(s)$ and simplifying, the expression above can be written as

$$f(s)A_K(s \rightarrow s') - f(s')A_K(s' \rightarrow s) = 0 \quad (52)$$

Now suppose that “ $s > s'$ ” (the other case is completely analogous) and substitute the appropriate acceptance probabilities in the above equation. Using eq. (25), we have to satisfy

$$f(s)(c_+ + c_- U(s \rightarrow s')) - f(s')(c_- + c_+ U(s' \rightarrow s)) = 0 \quad (53)$$

and, given that we have taken $U(s \rightarrow s') \rightarrow \rho'_m(\mathbf{r}')/\rho'_m(\mathbf{r})$ for $N \rightarrow \infty$ (see eq.(46)), eq. (53) is asymptotically satisfied.

Penalty

As mentioned in the text, when using the penalty method, D and χ^2 are stochastic variable distributed according to some unknown probability that describes the statistical noise associated to the calculation. The main idea in the method is that the acceptance probability associated to the exact Monte Carlo sampling of $P(s)$ can be expressed as the expectation value

$$A_P(s \rightarrow s') = \int_{-\infty}^{+\infty} dD \int_0^{+\infty} d\chi^2 \tilde{P}(D(s, s'), \chi^2(s, s')) a_P(D(s, s'), \chi^2(s, s')) \quad (54)$$

where $\tilde{P}(D, \chi^2)$, the so-called noise distribution, is the joint probability to realize a given value of the difference D and of the associated variance. The definition above is formal since the probability and the function a_P are not known. Under an appropriate hypothesis, however, this detailed balance expression can be used to obtain an explicit form for a_P . In fact, if the noise distribution is assumed to be symmetric under exchange of s and s' , the detailed balance condition can be expressed as

$$\begin{aligned} & \int_{-\infty}^{+\infty} dD \int_0^{+\infty} d\chi^2 \left\{ \tilde{P}(D(s, s'), \chi^2(s, s')) e^{-L(s)} \right. \\ & \times \left. [a_P(D(s, s'), \chi^2(s, s')) - e^{-\Delta} a_P(D(s, s'), \chi^2(s, s'))] \right\} = 0 \end{aligned} \quad (55)$$

Assuming also that $\tilde{P}$ can be written as the product of a Gaussian distribution (with mean value Δ) for D times a χ -square distribution for χ^2 , the relation above becomes an equation for $a_P(D(s, s'), \chi^2(s, s'))$. The solution of this equation is non trivial, but it has been shown in [31] that, for all Δ s, it is given by eq. (27) in section 3. Thus, since with this form of a_P the detailed balance relation is verified,

use of eq. (27) is a sufficient condition to generate a Markov chain that samples asymptotically $P(s)$.

Appendix II

Here we discuss how to extend our method to operators of the form $\hat{A} = \hat{p}^m \hat{A}_r \hat{p}^n$, with $\hat{A}_r$ diagonal in the coordinates representation. This will be done in two steps: First, assuming that all the required derivatives exist, we prove the identity

$$\hat{A}_r \hat{p}^n = \sum_{k=0}^n (i\hbar)^k \binom{n}{k} \hat{p}^{n-k} \hat{A}_r^{(k)} \quad (56)$$

where $\hat{A}_r^{(0)} = \hat{A}_r$ and $\hat{A}_r^{(k)}$ indicates the k -th derivative of $\hat{A}_r$ with respect to r . Second, we use this identity to write the matrix element $\langle r^0 + \frac{\Delta r^0}{2} | \hat{A} | r^0 - \frac{\Delta r^0}{2} \rangle$ (see eq.(4)) as a linear combination of terms proportional to the product of a function of r^0 times $\delta(\Delta r^0)$ (see eq. (6)). Using this result, the zero order symmetrized correlation function is written as a sum of terms, each calculable with the algorithm introduced in this paper.

Eq. (56) can be proved by induction. For $n = 0$, it is trivially $\hat{A}_r = \hat{A}_r$. Note also that, for $n = 1$ it reads

$$\hat{A}_r \hat{p} = \hat{p} \hat{A}_r + i\hbar \hat{A}_r^{(1)} \quad (57)$$

Assume now that the identity holds for n , then we can write

$$\hat{A}_r \hat{p}^{n+1} = \left\{ \hat{A}_r \hat{p}^n \right\} \hat{p} = \left\{ \sum_{k=0}^n (i\hbar)^k \binom{n}{k} \hat{p}^{n-k} \hat{A}_r^{(k)} \right\} \hat{p} \quad (58)$$

The product $\hat{A}_r^{(k)} \hat{p}$ at the end of the last equation above can be written using eq. (57), with $\hat{A}_r^{(k)} \rightarrow \hat{A}_r$. Substituting the result in the equation above, we get

$$\begin{aligned} \sum_{k=0}^n (i\hbar)^k \binom{n}{k} \hat{p}^{n-k} \hat{A}_r^{(k)} \hat{p} &= \sum_{k=0}^n (i\hbar)^k \binom{n}{k} \hat{p}^{n-k} (\hat{p} \hat{A}_r^{(k)} + i\hbar \hat{A}_r^{(k+1)}) \\ &= \sum_{k=0}^n (i\hbar)^k \binom{n}{k} \hat{p}^{n+1-k} \hat{A}_r^{(k)} + \sum_{k=0}^n (i\hbar)^{k+1} \binom{n}{k} \hat{p}^{n-k} \hat{A}_r^{(k+1)} \end{aligned} \quad (59)$$

Changing the summation index to $j = k + 1$ in the second sum of the last line above we can also write this last line as

$$\sum_{k=0}^n (i\hbar)^k \binom{n}{k} \hat{p}^{n+1-k} \hat{A}_r^{(k)} + \sum_{j=1}^{n+1} (i\hbar)^j \binom{n}{j-1} \hat{p}^{n+1-j} \hat{A}_r^{(j)}$$

“Resetting” $j = k$ in the second sum and isolating the term with $k = 0$ in the first sum and the term with $k = n + 1$ in the second, the line above can also be rearranged as

$$\hat{p}^{n+1} \hat{A}_r + \sum_{k=1}^n (i\hbar)^k \left[\binom{n}{k} + \binom{n}{k-1} \right] \hat{p}^{n+1-k} \hat{A}_r^{(k)} + (i\hbar)^{n+1} \hat{p}^{n+1} \hat{A}_r^{(n+1)} \quad (60)$$

We can now use the fact that $\binom{n+1}{0} = \binom{n+1}{n+1} = 1$ and Pascal's rule $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ to perform the last steps of the proof. Using these results in fact, the line above can be trivially written as

$$(i\hbar)^0 \binom{n}{0} \hat{p}^{n+1} \hat{A}_r + \sum_{k=1}^n (i\hbar)^k \left[\binom{n}{k} + \binom{n}{k-1} \right] \hat{p}^{n+1-k} \hat{A}_r^{(k)} + (i\hbar)^{n+1} \binom{n+1}{n+1} \hat{p}^{n+1-(n+1)} \hat{A}_r^{(n+1)} \quad (61)$$

or, writing everything as a single sum and remembering that we arrived at this expression via a sequence of equalities that started with eq.(58)

$$\hat{A}_r \hat{p}^{n+1} = \sum_{k=0}^{n+1} (i\hbar)^k \binom{n+1}{k} \hat{p}^{n+1-k} \hat{A}_r^{(k)} \quad (62)$$

which completes the proof by induction.

Let us now consider the matrix element

$$\begin{aligned} \langle r^0 + \frac{\Delta r^0}{2} | \hat{A} | r^0 - \frac{\Delta r^0}{2} \rangle &= \langle r^0 + \frac{\Delta r^0}{2} | \hat{p}^m \hat{A}_r \hat{p}^n | r^0 - \frac{\Delta r^0}{2} \rangle \\ &= \sum_{k=0}^n (i\hbar)^k \binom{n}{k} \langle r^0 + \frac{\Delta r^0}{2} | \hat{p}^{m+n-k} \hat{A}_r^{(k)} | r^0 - \frac{\Delta r^0}{2} \rangle \\ &= \sum_{k=0}^n (i\hbar)^k \binom{n}{k} A_r^{(k)} (r^0 - \frac{\Delta r^0}{2}) \langle r^0 + \frac{\Delta r^0}{2} | \hat{p}^{m+n-k} | r^0 - \frac{\Delta r^0}{2} \rangle \end{aligned} \quad (63)$$

Inserting a complete set of states in the momentum representation we can write the last line above as

$$\begin{aligned} &\sum_{k=0}^n (i\hbar)^k \binom{n}{k} A_r^{(k)} (r^0 - \frac{\Delta r^0}{2}) \int \frac{dp}{2\pi\hbar} p^{m+n-k} e^{-\frac{i}{\hbar} p \Delta r^0} \\ &= \sum_{k=0}^n (i\hbar)^k \binom{n}{k} A_r^{(k)} (r^0 - \frac{\Delta r^0}{2}) (-i\hbar)^{m+n-k} \frac{\partial^{(m+n-k)}}{\partial \Delta r^0 (m+n-k)} \int \frac{dp}{2\pi\hbar} e^{-\frac{i}{\hbar} p \Delta r^0} \\ &= \sum_{k=0}^n (-1)^{m+n-k} (i\hbar)^{m+n} \binom{n}{k} A_r^{(k)} (r^0 - \frac{\Delta r^0}{2}) \delta^{(m+n-k)}(\Delta r^0) \end{aligned} \quad (64)$$

With this expression for the matrix element, see eq. (4), $G^0(t, \beta)$ is given by the sum of n terms of the same structure. Let us consider the generic term “ k ”. Neglecting constant multiplicative factors we have

$$G_k^0(t, \beta) \propto \frac{\int d\Gamma e^{-\frac{i}{\hbar} p^1 \Delta r^{\nu-1}} \rho(\Gamma) A_r^{(k)} (r^0 - \frac{\Delta r^0}{2}) B_w(r_t, p_t) \delta^{(m+n-k)}(\Delta r^0)}{\int d\Gamma e^{-\frac{i}{\hbar} p^1 \Delta r^{\nu-1}} \rho(\Gamma) \delta(\Delta r^0)} \quad (65)$$

Using the definition of the derivative of the delta function (remembering that

$B_w(r_t, p_t)$ does not depend on Δr^0 , we get

$$G_k^0(t, \beta) \propto (-1)^{(m+n-k)} \frac{\int d\Gamma' e^{-\frac{i}{\hbar} p^1 \Delta r^{\nu-1}} B_w(r_t, p_t) \left[\frac{\partial^{(m+n-k)}}{\partial \Delta r^{0(m+n-k)}} \left(\rho(\Gamma) A_r^{(k)}(r^0 - \frac{\Delta r^0}{2}) \right) \right]_{\Delta r^0=0}}{\int d\Gamma e^{-\frac{i}{\hbar} p^1 \Delta r^{(\nu-1)}} \rho(\Gamma) \delta(\Delta r^0)} \quad (66)$$

Given the exponential form of $\rho(\Gamma)$ (see eq. (5)), we have, for any l and j ,

$$\left(\frac{\partial^l A_r^{(j)}(r^0 - \frac{\Delta r^0}{2}) \rho(\Gamma)}{\partial \Delta r^{0l}} \right)_{\Delta r^0=0} = \hat{Q} O_{l,j}(\Gamma') \hat{\rho}(\Gamma') \quad (67)$$

where Γ' is the set Γ minus the element Δr^0 and $O_{l,j}(\Gamma')$ is a function which can be analytically calculated and contains the derivatives up to order l of both the potential and the function $A_r^{(j)}$. Using the the result above in all terms of the expression for the symmetrized correlation function, we finally obtain

$$\begin{aligned} G_{AB}^0(t, \beta) &= \sum_{k=0}^n \left\{ (i\hbar)^{(m+n)} \binom{n}{k} \right. \\ &\quad \times \frac{\int d\Gamma' e^{-\frac{i}{\hbar} p^1 \Delta r^{\nu-1}} \hat{\rho}(\Gamma') O_{(m+n-k),k}(\Gamma') B_w(r_t, p_t)}{\int d\Gamma' e^{-\frac{i}{\hbar} p^1 \Delta r^{(\nu-1)}} \hat{\rho}(\Gamma')} \Big\} \\ &= \sum_{k=0}^n (i\hbar)^{(m+n)} \binom{n}{k} \langle O_{(m+n-k),k}(\Gamma') B_W(r_t, p_t) \rangle_{\mathcal{P}} \end{aligned} \quad (68)$$

Each term in the sum above can be computed via the generalized Monte Carlo scheme introduced in this work (in fact, a single Monte Carlo in which the n different averages are accumulated and then summed is sufficient), so the method we presented is in principle valid also for the more complex observables discussed here. Calculations of this kind will be more expensive than those of the test cases we discussed due to the structure of the observable. On the other hand, to close on an encouraging note, several interesting physical observables are defined with no, or very low order, powers of the momentum operator.

References

- [1] R. Hernandez and G.A. Voth, *Chemical Physics* **233**, 243 (1998).
- [2] H. Wang, X. Sun and W.H. Miller, *The Journal of Chemical Physics* **108**, 9726 (1998).
- [3] J.A. Poulsen, G. Nyman and P.J. Rossky, *The Journal of Chemical Physics* **119**, 12179 (2003).
- [4] Q. Shi and E. Geva, *The Journal of Physical Chemistry A* **107**, 9059 (2003).
- [5] M.S. Causo, G. Ciccotti, D. Montemayor, S. Bonella and D.F. Coker, *The Journal of Physical Chemistry B* **109**, 6855 (2005).
- [6] J. Liu and W.H. Miller, *The Journal of Chemical Physics* **128**, 144511 (2008).
- [7] T.D. Hone, J.A. Poulsen, P.J. Rossky and D.E. Manolopoulos, *The Journal of Physical Chemistry B* **112**, 294 (2008).
- [8] J.A. Poulsen, J. Scheers, G. Nyman and P. Rossky, *Physical Review B* **75**, 224505 (2007).
- [9] J.A. Poulsen, G. Nyman and P.J. Rossky, *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6709 (2005).
- [10] X. Sun, H. Wang and W.H. Miller, *The Journal of Chemical Physics* **109**, 7064 (1998).
- [11] S. Bonella and D.F. Coker, *The Journal of Chemical Physics* **122**, 194102 (2005).
- [12] P. Huo, S. Bonella, L. Chen and D.F. Coker, *Chemical Physics* **370**, 87 (2010).
- [13] R.P. Feynman and H. Kleinert, *Physical Review A* **34**, 5080 (1986).
- [14] J. Liu and W.H. Miller, *The Journal of Chemical Physics* **125**, 224104 (2006).
- [15] J. Liu and W.H. Miller, *The Journal of Chemical Physics* **127**, 114506 (2007).
- [16] D. Marinica, M. Gaigeot and D. Borgis, *Chemical Physics Letters* **423**, 390 (2006).
- [17] E. Jezek and N. Makri, *The Journal of Physical Chemistry A* **105**, 2851 (2001).
- [18] S. Bonella, M. Monteferrante, C. Pierleoni and G. Ciccotti, *The Journal of Chemical Physics* **133**, 164104 (2010).
- [19] S. Bonella, M. Monteferrante, C. Pierleoni and G. Ciccotti, *The Journal of Chemical Physics* **133**, 164105 (2010).
- [20] V. Filinov, *Molecular Physics* **88**, 1517 (1996).
- [21] V. Filinov, *Molecular Physics* **88**, 1529 (1996).
- [22] W. Miller, S. Schwartz and J. Tromp, *The Journal of Chemical Physics* **79**, 4888 (1983).
- [23] V. Jadhao and N. Makri, *The Journal of Chemical Physics* **129**, 161102 (2009).
- [24] J. Poulsen, H. Li and G. Nyman, *The Journal of Chemical Physics* **131**, 024117 (2009).
- [25] G. Krilov, E. Sim and B. Berne, *Chemical Physics* **268**, 21 (2001).
- [26] G. Krilov, E. Sim and B. Berne, *The Journal of Chemical Physics* **114**, 1075 (2001).
- [27] N. Chakrabarti, T. Carrington and B. Roux, *Chemical Physics Letters* **293**, 209 (1998).
- [28] J. Kegerreis, A. Nakayama and N. Makri, *The Journal of Chemical Physics* **128**, 184509 (2008).
- [29] H. Kleinert, *Path Integrals in Quantum Mechanics, Statics, Polymer Physics and Financial Markets* (World Scientific, Singapore, 2004).
- [30] A. Kennedy and J. Kuti, *Physical Review Letters* **54**, 2473 (1985).
- [31] D.M. Ceperley and M. Dewing, *The Journal of Chemical Physics* **110**, 9812 (1999).
- [32] D.V. Lindley, *Introduction to probability and statistics* (Cambridge University Press, Cambridge, 1965).
- [33] M. Sprik, M. Klein and D. Chandler, *Physical Review B* **31**, 4234 (1985).
- [34] D.M. Ceperley, *Review of modern physics* **67**, 279 (1995).
- [35] V. Jadhao and N. Makri, *The Journal of Chemical Physics* **129**, 161102 (2008).

