

Determination of effective pair interactions from the structure factor

N. G. Almarza, E. Lomba, and D. Molina

Instituto de Química Física Rocasolano (CSIC), C/Serrano 119, E-28006 Madrid, Spain

(Received 20 February 2004; published 13 August 2004)

In this work we present an efficient procedure to evaluate effective pair potentials, compatible with “experimental” structure factors, using a Monte Carlo simulation scheme. The procedure does not require the use of inverse Fourier transforms and is robust and rapidly convergent. As a test case the structure factor of liquid Selenium obtained from a Tight-Binding Molecular Dynamics simulation is inverted to obtain an effective pair potential and, as a by-product, the pair distribution function. The inversion procedure yields a pair structure in perfect agreement with the original molecular dynamics calculations and the analysis of the triplet structure and the dynamics also illustrates the limitations of the use of pair potentials in the description of liquids with strongly directional bonding, such as the covalent liquid Selenium.

DOI: 10.1103/PhysRevE.70.021203

PACS number(s): 61.20.Ja, 05.20.Gg, 02.70.Uu

I. INTRODUCTION

The determination of the Hamiltonian of a given system from its microscopic structure—what is known as the inverse problem—remains one of the crucial challenges in the physics of disordered condensed matter. In the last two decades numerous approaches have been formulated, most of them based on the prior knowledge of the microscopic structure in terms of the pair distribution function (PDF) $g(r)$ [1–7]. The most efficient procedures resort to some sort of simulation method [5–7] to obtain an effective pair potential compatible with the PDF used as input. This quantity, however, is not directly accessible from the experiment. One must resort to the evaluation of inverse Fourier transforms (i.e., $k \rightarrow r$) of the structure factors $S(k)$ obtained from x-ray or neutron diffraction experiments. This inverse Fourier transformation is plagued with numerical inaccuracies, mainly stemming from the limited k -range for which the scattering data are available. The extension of Soper’s method [6] to incorporate structure factors as input [8], although somewhat different in its formulation, also requires the evaluation of an inverse Fourier transform. On the other hand, the original Reverse Monte Carlo (RMC) technique [9], even if it is perfectly suitable to generate microscopic configurations compatible with the input $S(k)$, is not a solution of the inverse problem. Moreover, it seems to be too prone to predict structures with an excess of entropy [10]. This will not be the case in methods based on the use of effective potentials, in which the energy is constrained [8].

Quite recently Toth [11,12] has proposed an extension of the method of Lyubartsev and Laaksonen [7], which uses as input data $S(k)$, involving only direct Fourier transforms (i.e., $r \rightarrow k$). In this procedure just the PDF of the simulated sample has to be Fourier transformed, and this can be reliably done if the sample size is sufficiently large. Toth’s method in its last version [12] seems powerful and efficient, but in each refinement step it requires the numerical solution of a set of nonlinear equations with n -unknowns, n being the number of parameters that determine the potential—if its functional form is known—or the number of data points for which the tabulated effective pair potential is to be evaluated.

This latter case, which is the most common situation, constitutes a highly nontrivial numerical problem.

The purpose of this paper is to show that a robust, simple, and computationally inexpensive algorithm can be constructed on the basis of a recently proposed method for the determination of effective pair potentials from PDFs [13]. In order to illustrate the capabilities of this new algorithm, it will be put to an stringent test, the inversion of the structure factor of liquid Selenium computed from a Tight-Binding Molecular Dynamics (TBMD) simulation [14]. It will be shown that the proposed method can reliably invert the structure factor even with truncated k -range, rendering a pair structure in perfect accordance with that of the TBMD simulation. The limitations of the use of pair potentials in systems with strong directional bonding like this, will in turn be made apparent when calculating other properties, such as the vibrational spectrum or the three-particle distribution function.

In practice we will deal with the problem of the simultaneous determination of the PDF and the effective pair potential for a given system, whose structure factor $S(k)$ is known. For the sake of simplicity, we will focus here on a monoatomic system. Generalizations to molecular systems are possible provided that the different partial structure factors are known. The structure factor [15] of a monoatomic (and isotropic) system is given by

$$S(k) = 1 + 4\pi\rho \int_0^\infty dr r^2 h(r) \frac{\sin(kr)}{kr}, \quad (1)$$

where $h(r) \equiv g(r) - 1$, r stands for the distance and ρ is the density. Let us assume that we know the experimental static structure factor of a certain system at given conditions of density and absolute temperature T ; our aim will then be to determine an effective pair potential $V(r)$ that accounts for the pair structure of the system under consideration. In other words, if we perform a computer simulation of a condensed phase whose interactions can be described by $V(r)$ at the same thermodynamic conditions, then the calculated structure factor $S_{sim}(k)$ must be equal (within error bars) to the experimental result. In what follows, how to achieve this goal will be explained in detail.

II. NUMERICAL PROCEDURE

Let us suppose that we have a certain number of experimental data for the structure factor (together with a corresponding estimation of their error bars): $[k_i, S(k_i), \Delta S(k_i)]$ with $i=1, 2, \dots, N_k$, or in shorter notation $(k_i, S_i, \Delta S_i)$, for a system with density ρ . We will represent the effective interaction potential as a 3D Fourier sum of the form

$$\beta V(r) = \sum_{i=1}^{N_k} u_i \frac{\sin(k_i r)}{k_i r}, \quad (2)$$

where $\beta=1/k_B T$ as usual. Thus, our inversion problem consists in determining the values of the coefficients u_i , ($i=1, 2, \dots, N_k$) such that (2) leads to a simulated $S_{sim}(k)$ consistent with the experimental values. In practice, the effective interaction $\beta V(r)$ will be handled in terms of discrete values calculated at $r_i=i\Delta r$, which will then be interpolated for arbitrary r , and will be truncated according to the simulation box size.

The inversion procedure makes use of the Metropolis Monte Carlo method [16,17] and is divided in several stages. At each stage we will be performing changes in the coefficients u_i in order to achieve the matching between the input structure factor and that obtained within the simulation. Such changes will be smaller as we advance in successive stages. The length of the simulation run on each stage will depend on the fulfillment of convergence criteria applied to the simulated structure factor when compared to the input $S(k)$. After a cycle (N translation attempts, being N the number of particles) of simulation we compute an instantaneous structure factor $S_{inst}(k)$ by performing the Fourier transformation of the instantaneous PDF, $g_{inst}(r)$. These data are accumulated into the corresponding functions of the stage. The $S_{inst}(k_i)$ are determined as

$$S_{inst}(k_i) = 1 + 4\pi\rho \int_0^{r_c} dr r^2 [g_{inst}(r) - 1] \frac{\sin(k_i r)}{k_i r}, \quad (3)$$

where r_c depends on the system size. If the sample size is reasonably large one can rely on the use of a direct Fourier transformation like the one of Eq. (3). For particularly involved systems (some liquid metals exhibiting a marked long-range structure in the PDF), it is possible resort to an extension method of the type devised by Verlet [18], although with present-day computers one can simply increase the sample size. For the case studied here, we found no significant differences between the extension procedure and the direct Fourier transformation. Now, the new values of the coefficients u_i are given [13] by

$$u_i^{new} = u_i^{old} + \lambda_s \frac{S_{inst}(k_i) - S_i}{\Delta S_i} \langle \Delta S \rangle, \quad (4)$$

where

$$\langle \Delta S \rangle = \frac{1}{N_k} \sum_{i=1}^{N_k} \Delta S_i, \quad (5)$$

and λ_s is a parameter that depends on the stage, s , of the calculation. With the new values of the coefficients u_i it is

possible to build up the new effective potential, by means of Eq. (2). After each block of a given number of simulation cycles, the fulfillment of the following convergence criterion is checked

$$\frac{1}{N_k} \sum_{i=1}^{N_k} \left[\frac{S_s(k_i) - S_i}{\Delta S_i} \right]^2 < \eta, \quad (6)$$

where the value of the convergence parameter η is predetermined (with values typically ≤ 1 , see Sec. III for actual values), and $S_s(k)$ represents the accumulated structure factor through the whole stage. If Eq. (6) is fulfilled, the procedure initiates the subsequent stage, which starts with the current values of the coefficients u_i . On the new stage, $s+1$, the parameter λ is given by

$$\lambda_{s+1} = \alpha \lambda_s = \alpha^s \lambda_1, \quad (7)$$

with $0 < \alpha < 1$.

As starting point we have chosen $u_i=0$, for all i . The initial choice corresponds to a system of noninteracting particles. Other initial guesses are possible, and it is likely that initial solutions closer to the actual effective potential will speed up the convergence of the method. As the simulation evolves, an effective pair potential is expected to develop. Such a situation makes it convenient to introduce a mechanism to control the maximum displacement parameter δx_{max} of the single particle MC moves [16], in order to keep a reasonable acceptance ratio of trial configurations. This idea can be put into practice by computing on each block of the procedure the fraction of accepted moves X_A . After each block we can slightly modify the value of δx_{max} in order to push the system toward a certain predetermined target acceptance ratio (e.g., 1/2), for instance using

$$\ln \delta x_{max}^{new} = \ln \delta x_{max}^{old} + \tau_s (X_A - 1/2); \quad \tau_s \geq 0, \quad (8)$$

where the value of the multiplier τ depends on the simulation stage.

It is known that the change of the parameter δx_{max} during the averaging configurations of a simulation can break the detailed balance fulfillment [17]. In the present case such a circumstance does not seem an important problem (at least in the first stages of the procedure) because the change of the running effective pair potential also implies the violation of detailed balance. Nevertheless, we expect to approach equilibriumlike conditions in successive stages, therefore, it seems sensible either to reduce the value of τ when starting new stages or, eventually, eliminate the δx_{max} updating device after a number of initial stages. We have chosen the first possibility by using

$$\tau_s = \alpha^{s-1} \tau_1. \quad (9)$$

Using this scheme and provided that the simulation results converge at the different stages (which we expect to happen if the effective pair potential exists, see [13]), the detailed balance will become practically accomplished in the last stages of the procedure, since the changes in both u_i coefficients and δx_{max} along each stage will become vanishingly small.

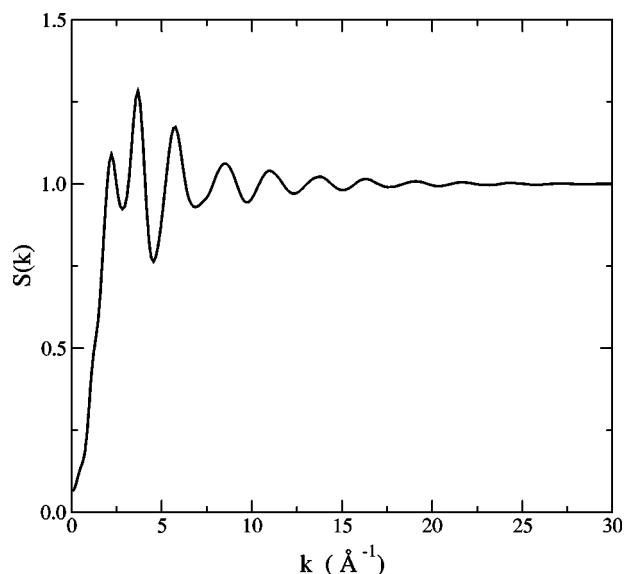


FIG. 1. Structure factor of liquid selenium evaluated from Fourier transformation of the results of $g(r)$ obtained in the TBMD simulation.

III. APPLICATION

The “experimental” structure factor has been calculated by means of a TBMD simulation carried out on the micro-canonical ensemble, with 416 particles at $\rho=0.0295 \text{ \AA}^{-3}$, and a time step of 1 fs. For the present model this state corresponds to a high-density liquid at a pressure of 5.1 ± 1.4 kbar, well away from the triple point. The system was equilibrated for 10 ps after melting the initial configuration, and the production run included 20 ps more. The average temperature was 570 K. From this calculation we evaluated the pair distribution function up to a distance 12.1 \AA . At this distance $g(r) \approx 1$ within experimental error, and one may expect the corresponding structure factor to be accurate enough. This latter quantity is tabulated in k -space with a grid $\Delta k = 0.12272 \text{ \AA}^{-1}$. Block averages were used to estimate the errors in both $g(r)$ and $S(k)$. This function is plotted in Fig. 1 and the corresponding pair distribution is represented in Fig. 2. We note in passing that the high and narrow first peak of $g(r)$ simply results from the presence of covalent bonds in the sample.

Using the TBMD results of the PDF we have applied the inversion method proposed in Ref. [13] (Case A) (with convergence parameter $\eta=0.10$). Then we have inverted the structure factor following the scheme proposed in this work, using a number of stages $N_s=30$, $\lambda_1=10.0$, $\alpha=0.75$, and $\eta=0.10$. In order to analyze the effect of truncating $S(k)$ we have run the procedure using three different values of N_k : Case B: $N_k=326$, $k_{max} \approx 40.0 \text{ \AA}^{-1}$, Case C: $N_k=122$, $k_{max} \approx 15.0 \text{ \AA}^{-1}$, and Case D: $N_k=81$, $k_{max} \approx 10.0 \text{ \AA}^{-1}$. In all cases the number of particles was $N=500$.

In Fig. 3 we show the results of the effective potentials evaluated using the four alternatives indicated above. The first conclusion that can be extracted from Fig. 3 is that the main differences between the results are located at short distances. Such differences do not appear to be relevant, since

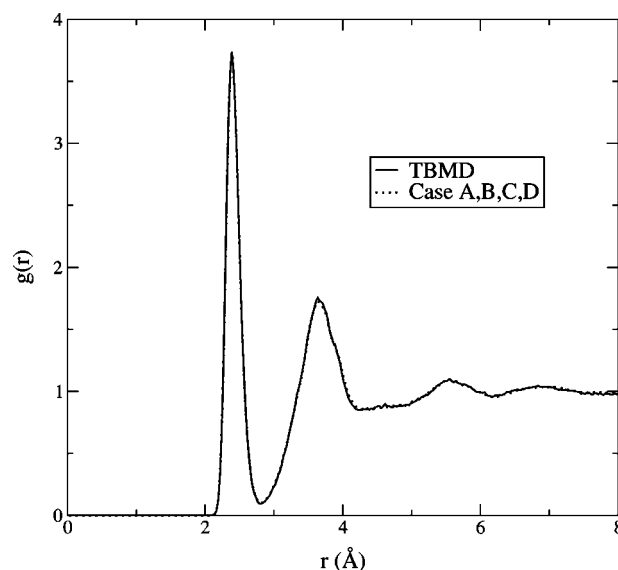


FIG. 2. PDF for liquid Se. The solid curve represents the TBMD results, the dotted line corresponds to models A–D (see Sec. III and Fig. 3 for details of the models), which are indistinguishable at the scale of the figure.

the interaction potential is very high in this range and no important contributions to the structure of the fluid are expected. On the other hand, it is remarkable the excellent agreement between the different results for $r \geq 2.1 \text{ \AA}$. In principle, the agreement between Cases A and B, could be expected since in both cases the “experimental” functions are well represented taking into account their respective cutoffs. However, for Cases C and D, the values of k_{max} correspond to a region of wave numbers where considerable structure is still present (see Fig. 1). Nonetheless this lack of information on the tails of $S(k)$ seems to have little influence on the relevant range of the effective potential.

The small differences between the results of the two types of inversion (see Fig. 3) for large distances are probably due to the boundary conditions. In the inversion of $g(r)$ we set a cutoff distance of $r_c \approx 12.1 \text{ \AA}$ (according to the size of the simulation box of the TBMD run), whereas its value in the inversions using $S(k)$ was $r_c \approx 12.82 \text{ \AA}$. Note that in all cases the resulting PDF (see Fig. 2) is practically indistinguishable from the TBMD $g(r)$.

IV. QUALITY OF THE EFFECTIVE POTENTIALS

In order to analyze the ability of a pair interaction model to give a correct account of various properties of the TB model for liquid Se, we have run Molecular Dynamics simulations using the effective potentials extracted from Cases A and B, and compared the results with those of the TBMD calculation. The extracted effective potentials are smoothed with a least-squares piecewise cubic spline.

As expected, the radial distribution functions coming from the pair interaction model were (within statistical errors) equal to that evaluated in the TBMD calculation. Now, if one goes beyond the pair structure, things start to look different. In Fig. 4 we have plotted the three particle distri-

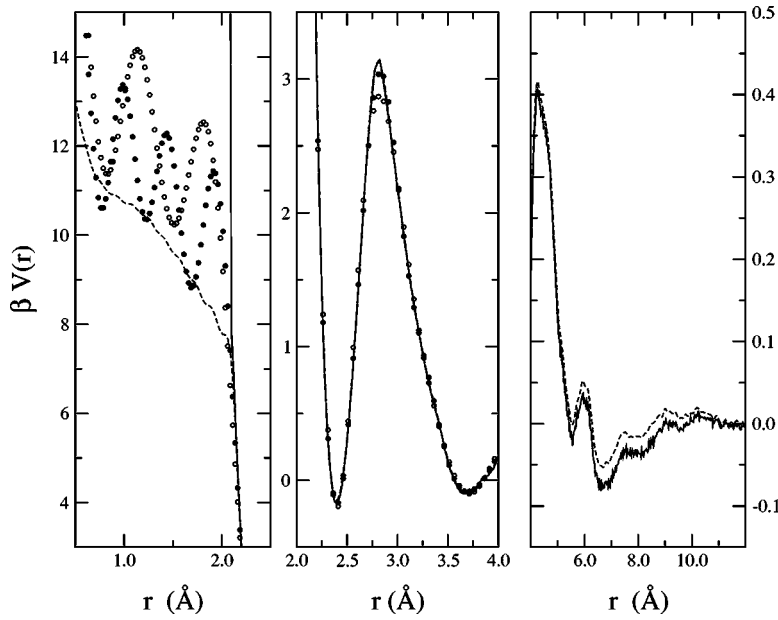


FIG. 3. Effective pair potentials evaluated using inversion techniques. Case A {inversion of $g(r)$ following the method of [13]}; continuous line. Case B ($k_{max}=40 \text{ \AA}$): dashed line. Case C ($k_{max}=15 \text{ \AA}$): filled circles. Case D ($k_{max}=10 \text{ \AA}$): white circles. In the right frame, only Cases A and B are represented (since Cases C and D are indistinguishable from Case B at the scale of the figure).

bution function $g^{(3)}(r, s, \theta)$, calculated for an isosceles configuration with $r=s=2.48 \text{ \AA}$. This distance corresponds to the first peak in $g(r)$. We observe that, whereas the TBMD triplet distribution exhibits a single maximum at 100° , which is the proper bond angle for the experimental liquid Se, the pair potential presents an additional very marked maximum at 60° . This results from the fact that a pair potential cannot prevent the formation of triplets of Se atoms in equilateral triangular configurations. These Se_3 clusters are absent from the TBMD and the experimental results. If we now look at the vibrational density of states depicted in Fig. 5, we observe that the overall shape of the curve is reasonably reproduced by the pair potential model, but the position of the stretching band at 30 meV is underestimated, the bond bending band between 8–15 meV is less broad and finally $z(0)$ is much overestimated. This latter feature simply reflects the

fact that the diffusion constant in the pair potential model is much higher ($4.8 \times 10^{-9} \text{ m}^2/\text{s}$) than that of the TB model ($1.2 \times 10^{-9} \text{ m}^2/\text{s}$). What is clear from these results is that the atoms are less tightly bound in the pair potential model, by which the diffusion constant is much higher and the bond stretching band less energetic. A coordination analysis indicates that whereas the TB model leads to an overall domination of doubly and singly coordinated atoms, isolated atoms and threefold coordination are also important in the pair potential models. Finally, it is worth pointing out that Cases A and B (i.e., those obtained from the $g(r)$ and the $S(k)$ inversion, respectively) lead to identical vibrational dynamics as can be seen in the coincident $z(E)$ curves in Fig. 5. This is not surprising since the bond dynamics is essentially determined by the short- and medium-range structure of the potential and in Fig. 3 can readily be seen that Cases A and B are hardly distinguishable in the region $2 \text{ \AA} - 6 \text{ \AA}$.

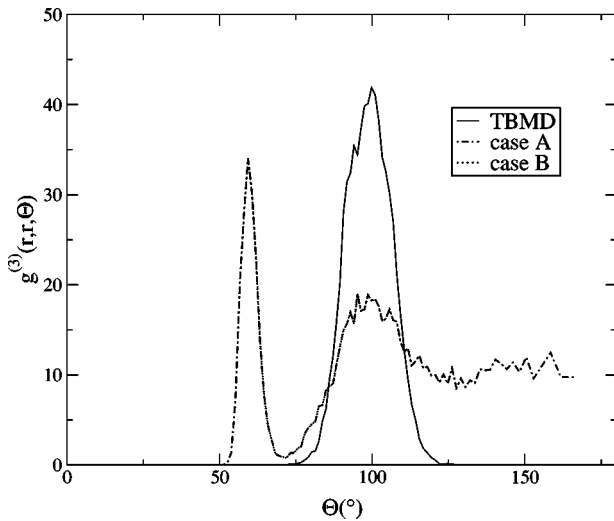


FIG. 4. Three-particle distribution function $g^{(3)}(r, s, \theta)$ calculated for $r=s=2.48 \text{ \AA}$, nearest-neighbor distance in liquid Se.

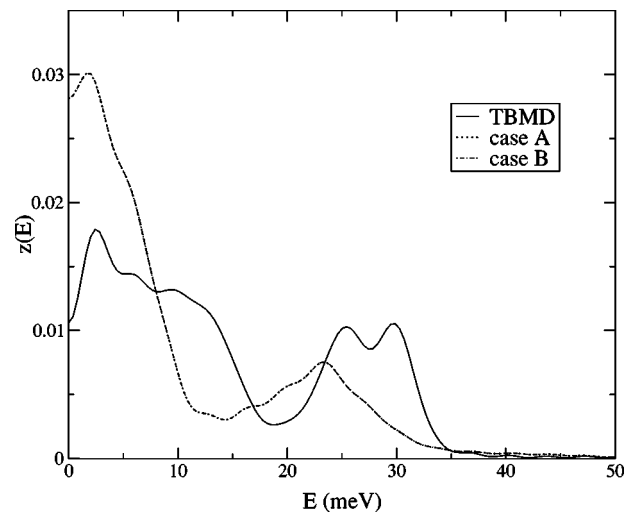


FIG. 5. Vibrational density of states for liquid Se.

V. CONCLUSIONS

In summary, we have presented an efficient and simple procedure to extract pair potentials from structure factors, which incidentally determines the pair distribution function without resorting to inverse Fourier transforms. The procedure can easily incorporate additional input information, such as thermodynamics (see Ref. [13]), in order to improve the quality of the potential. On the other hand, the test case chosen, liquid Se, illustrates the limitations intrinsic to the use of pair potentials when one tries to model properties that go beyond the pair structure in systems characterized by strong directional bonding. In a forthcoming work it will be shown that the method can be generalized to simultaneously

extract effective pair and triplet potentials, if information on the three-body pair distribution is incorporated as input. This will pave the way to the development of a systematic procedure to reduce complex interactions, like those present in biomolecules, which cannot be represented by mere effective pair potentials.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the Dirección General de Investigación Científica under Grants No. MAT2002-04540-C05-C03/05 (N.G.A.) and No. BFM2001-1017-C03-01 (E.L.).

-
- [1] R.L. Henderson, *Phys. Lett.* **49A**, 197 (1974).
[2] M.D. Johnson, P. Hutchinson, and N.H. March, *Proc. R. Soc. London, Ser. A* **282**, 283 (1964).
[3] W.S. Howells and J.E. Enderby, *J. Phys. C* **5**, 1277 (1972).
[4] G. Kahl, B. Bildstein, and Y. Rosenfeld, *Phys. Rev. E* **54**, 5391 (1996).
[5] D. Levesque, J.J. Weis, and L. Reatto, *Phys. Rev. Lett.* **54**, 451 (1985).
[6] A.K. Soper, *Chem. Phys.* **202**, 295 (1996).
[7] A.P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* **52**, 3730 (1995).
[8] T. Yamaguchi, K. Hidaka, and A.K. Soper, *Mol. Phys.* **96**, 1159 (1999).
[9] R.L. McGreevy and L. Pusztai, *Mol. Simul.* **1**, 369 (1988).
[10] P. Jedlovsky, I. Bako, G. Palinkas, and A.K. Soper, *J. Chem. Phys.* **105**, 245 (1996).
[11] G. Toth, *J. Chem. Phys.* **115**, 4770 (2001).
[12] G. Toth, *J. Chem. Phys.* **118**, 3949 (2003).
[13] N.G. Almarza and E. Lomba, *Phys. Rev. E* **68**, 011202 (2003).
[14] D. Molina, E. Lomba, and G. Kahl, *Phys. Rev. B* **60**, 6372 (1999).
[15] J.P. Hansen and I.R. McDonald, *Theory of Simple Liquids* (Academic, Oxford, 1986), pp. 97–99.
[16] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1986), pp. 118–123.
[17] D. Frenkel D. and B. Smit, *Understanding Molecular Simulation* (Academic, New York, 1996), pp. 19–52.
[18] L. Verlet, *Phys. Rev.* **165**, 202 (1968).