

Simplification of the molecular dynamics that preserves thermodynamics

Sebastián Carrasco*

Departamento de Física, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile 7800024

José Rogan and Juan Alejandro Valdivia

*Departamento de Física, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile 7800024
and Centro para la Nanociencia y la Nanotecnología, CEDENNA, Chile*

(Received 1 June 2018; revised manuscript received 23 July 2018; published 10 December 2018)

A systematic algorithm to reduce the potential energy surface is presented, which is based on successive inversions of virial data, so that the molecular dynamics is capable of preserving the thermodynamics of the system. The algorithm is tested by reducing the potential energy surface of diatomic molecules that interact through a non-spherically symmetric potential, into an effective potential which is spherically symmetric, showing that all rotational degrees of freedom of the molecules can be disregarded from the thermodynamic point of view, as long as the virial expansion is a good description of the system. Furthermore, our analysis allows us to understand how we can make use of the non-uniqueness of the solution of an inverse problem to generate simplified models.

DOI: [10.1103/PhysRevE.98.063308](https://doi.org/10.1103/PhysRevE.98.063308)**I. INTRODUCTION**

Much work and several strategies have been developed to establish a good parametrization of the potential-energy surface (PES) for interacting atoms or molecules. The interaction energies are then used to make quantitative predictions about dynamical processes like energy transfer, scattering, and chemical reactions [1]. However, in such calculations, computational time grows quite fast with the degrees of freedom and the number of particles, eventually making the system quite expensive computationally. One possible way to avoid these constraints is to consider that the thermodynamic behavior of these systems usually do not depend on every detail stored in the PES, so that a simpler effective PES could be found. This idea is supported by many coarse-grain (CG) models that reproduce many properties of the systems under consideration. These CG models have become widely used as they are applied in simplified water models [2–12], solvent free models [13–20], and models that include chemical specificity [21–23]. The development of the form of the functional used in these models is normally based on physical intuition. Of course, there are no guarantees that these potential descriptions, or reductions, provide a good representation of the molecular dynamics of the interacting particles or molecules, i.e., that the potential, although simplified, it reproduces the correct thermodynamics.

Hence, in this article we present a systematic method to reconstruct an effective PES parametrization in an unbiased way, with fewer degrees of freedom than the starting one, but with the same thermodynamics properties. This opens the door to a new way to approach the problem of molecular interactions, making possible a simplification of the problem both

theoretical and in practical applications. To do so, we propose to reconstruct an effective PES from the virial coefficients at different temperatures. As a test of our method, we present the reduction of the potential energy surface of a non-spherically symmetric variant of a Lennard-Jones diatomic potential into an effective potential which is spherically symmetric, showing that all rotational degrees of freedom of the molecules are negligible from the thermodynamic point of view at low enough density. We must say that this is not the only way to solve this problem at hand, but certainly is an alternative strategy to established approaches like iterative Boltzmann inversion [24], force-matching [25,26], inverse Monte Carlo [27], and relative entropy [28].

We choose the inversion of the virial coefficients for three reasons. First, they contain all the thermodynamic properties of the system, namely the state equations for pressure and energy, overcoming some limitations of other methods in representing thermodynamic properties like pressure and compressibility [29,30]. Second, the PES can be parametrized sequentially, solving for a two-body potential given the second virial data. After that we can solve for a three-body potential given the third virial data, and so on. Finally, the procedures to perform the inversion have been successfully developed before [31,32].

In contrast to the calculation of the virial coefficients given the PES, the reconstruction of the PES given the virial coefficients is an ill-posed problem [31], where existence, uniqueness, and continuity are not all fulfilled. We will see that the non-uniqueness of the solution of this inverse problem is fundamental to the generation of simplified models. Furthermore, the inversion of the second virial coefficient can be attempted in a number of ways. One way is to first introduce a specific expression for a potential and then try to fit the parameters of the short- and long-range two-body potential separately [31], which usually are represented in different

*sebastian.carrasco@ug.uchile.cl

temperature ranges of the virial coefficients. Another way is to use a nonlinear inversion procedure, such as was proposed by Sebastião *et al.* [32] using a neural network (NN). Hence, here we will adapt the non-unique nonlinear inversion procedure based on NNs to construct a simplified model for the particle interaction (reduction of the PES) that is consistent with the statistical mechanics of the original system, and therefore, the molecular dynamics. This method does not involve a specific form of the potential and so it allows to solve for the short and long range of the potential at the same time. In this work, we will use a NN for the inversion of the second and third virial data to generate a reduced potential. Hence, the NN give us an unbiased approach to the problem, saving us from choosing a functional form.

Our approach can be related to the relative entropy minimization approach, where a CG model is constructed by imposing that the probability of finding a microstate in the ensemble is the same for the model and the target system [28]. In our approach, such condition is imposed sequentially, similar to a cluster expansion in statistical mechanics, because virial coefficients are closely related to the canonical partition function. And so, it constitutes an alternative to the use of potentials and mean force approaches, like gaussian approximation potentials, where the mean forces are constructed sequentially in a cluster expansion manner [25,26].

By construction, the convergence of our approach depends on the convergence of the virial expansion, and so is applicable for small densities only. It is possible to generalize our approach by expanding around another density ρ_o of interest, e.g., $\rho_o \approx 1$, as shown in Ref. [33]. We plan to work on this interesting idea in a future manuscript. In a similar way, it can be generalized for multicomponent systems, using a multicomponent virial expansion as it is presented in Ref. [34].

II. NONLINEAR INVERSION PROCEDURE

The virial equation of state is the standard way to represent the deviation of the pressure p from a non-interacting behavior as an infinite power series in the number density ρ , namely [35],

$$\frac{p}{k_B T} = \rho + B(T)\rho^2 + C(T)\rho^3 + \mathcal{O}(\rho^4), \quad (1)$$

where k_B is the Boltzmann constant and T the temperature. The first virial coefficient is identically 1 and represents the case of non-interacting particles, B is the second virial coefficient and C the third virial coefficient. Similarly, the internal energy U can be expanded as

$$U = U_0 - Nk_B T^2 \frac{\partial B(T)}{\partial T} \rho - Nk_B T^2 \frac{\partial C(T)}{\partial T} \rho^2 + \mathcal{O}(\rho^3), \quad (2)$$

where U_0 is the internal energy in the non-interacting limit, which can be adjusted adding internal degrees of freedom, and N the number of molecules. Note that all the thermodynamic properties can be derived from these two equations, and consequently from the virial coefficients.

For second virial inversion, we start considering a NN with the following architecture: one perceptron on the input layer, one hidden-layer with n perceptrons, and one perceptron on

the output layer. The hidden layer uses a tanh as an activation function and the output layer considers linear function. For the sake of simplicity, we can picture the NN as a function with many parameters, namely,

$$N_2(r) = b_0 + \sum_{j=1}^n w_{2,j} \tanh(w_{1,j}r + b_j), \quad (3)$$

where b_j and w_{ij} are the biases and weights of the NN to adjust. In order to have a reasonable representation of the potential with physical sense, we choose b_0 so that $N(r) = 0$ when $r \rightarrow \infty$, so this is a constrained NN.

We first construct the second virial coefficient as a function of temperature of a non-spherically symmetric variant of the Lennard-Jones potential for diatomic molecules, that considers the structure of the diatomic molecule as if all atoms in the diatomic molecule interact by standard Lennard-Jones potential with atoms of other diatomic molecules. This adds an angular interaction of a dipole type given by $u(r, \Omega_1, \Omega_2)$, where r is the distance between the diatomic molecules, and Ω_1 and Ω_2 are the solid angles representing the absolute orientation of each diatomic molecule.

The second virial coefficient can be calculated from [35]

$$B(T) = -\frac{1}{2} \left\langle \int f_{12}(u) d\vec{r}_{12} \right\rangle_{\Omega_1, \Omega_2}. \quad (4)$$

Here, T is the non-dimensional temperature, $u(r_{12}, \Omega_1, \Omega_2)$ the potential between two molecules and

$$f_{12}(u) = \exp\left(-\frac{u(r_{12}, \Omega_1, \Omega_2)}{k_B T}\right) - 1, \quad (5)$$

the Mayer- f function for the potential u .

Our goal is to make a spherically symmetric NN that reproduces this non-spherically symmetric potential, but reproduces all of its thermodynamic properties for low enough density. To do so, let us define the cost function as

$$S_2 = \frac{1}{2M} \sum_{j=0}^M [B^{\text{obj}}(T_j) - B^{\text{NN}}]^2. \quad (6)$$

Note that the cost function is zero precisely when the objective second virial coefficient B^{obj} is the same for the potential produced by the NN and the original potential over all desired temperatures. So, if we perform a minimization we can find the correct set of weights and biases to construct the potential provided by the NN which will have the same second virial coefficient than the original potential $u(r, \Omega_1, \Omega_2)$, but without considering the angular interaction. In this work, we use the gradient descent algorithm, with several random seeds, as a minimizer.

To compute the second virial of a given potential we use the Gauss-Laguerre quadrature, because it is a fast way to compute the integral, but we stretch the function by an α factor to improve convergence, namely,

$$\int_0^\infty f_{12}(N_2) r_{12}^2 dr_{12} \approx \sum_{i=1}^N \frac{\omega_i e^{x_i} x_i^2}{\alpha^3} f_{\frac{x_i}{\alpha}}(N_2), \quad (7)$$

where x_i is the i th root of the Laguerre polynomial and w_i the respective weight.

For third virial inversion, we start considering a NN with the following architecture: three perceptrons on the input layer, to represent the three input distances; one hidden-layer with n perceptrons; and one perceptron on the output layer. We use \tanh as the activation function in the hidden layer and a linear function in the output layer. To ensure the particle exchange symmetry we sum over the set of all permutations (P) for the input layers, namely,

$$N_3(r_1, r_2, r_3) = b_0 + \sum_{j=1}^n \sum_{a,b,c \in P_3} \omega_{2,j} \tanh(|\omega_{1,j}^1| r_a + |\omega_{1,j}^2| r_b + |\omega_{1,j}^3| r_c + b_j), \quad (8)$$

where b_j and w_{ij} are the biases and weights of the NN to adjust. In order to have a representation of the potential with more physical sense we choose b_0 so that $N_3 = 0$ if any distance goes to infinity. This is the reason to have only positive weights, so we obtain a constrained and symmetrized NN.

The third virial coefficient is given by [35]

$$C(T) = C_{\text{add}} + \Delta C_{\text{nadd}}, \quad (9)$$

where the contribution of the pairwise additive terms is given by

$$C_{\text{add}}(T) = -\frac{1}{3} \left\langle \iint f_{12}(u_2) f_{13}(u_2) f_{23}(u_2) d\vec{r}_{12} d\vec{r}_{23} \right\rangle_{\Omega_1 \Omega_2 \Omega_3} \quad (10)$$

and the contribution of the three-body terms is given by

$$C_{\text{nadd}} = -\frac{1}{3} \left\langle \iint (e^{-u_3/T} - 1) \times e^{-[u_2(r_{12})+u_2(r_{13})+u_2(r_{23})]/T} d\vec{r}_{12} d\vec{r}_{23} \right\rangle_{\Omega_1 \Omega_2 \Omega_3}, \quad (11)$$

where u_3 is the three-body non-additive contribution.

In analogy with the previous second virial case, we define the cost function by

$$S_3 = \frac{1}{2M} \sum_{j=0}^M [C^{\text{obj}}(T_j) - C^{\text{NN}}(T_j)]^2. \quad (12)$$

In every computation of C we use the Gauss-Laguerre quadrature, including the averaging over all orientations. So, if we perform a minimization over all weights and biases of N_3 using N_2 as the corresponding pairwise potential, we will obtain a pair of effective potentials, namely, N_2 and N_3 , that have the same second and third virial coefficient than the original potentials u_2 and u_3 , but with no rotational dependence.

It is straightforward to extend this method for inversion to higher virial coefficients, which will allow to obtain four-body potentials and so on. However, we decided to truncate the expansion, which could be technically extended to any order multibody potential [36], to the three-body potential not only for simplicity, but also because it is flexible enough to provide a good account for the interactions on systems as complex as water [10], or a model for self-assembly of amphiphiles in vesicles [13], etc.

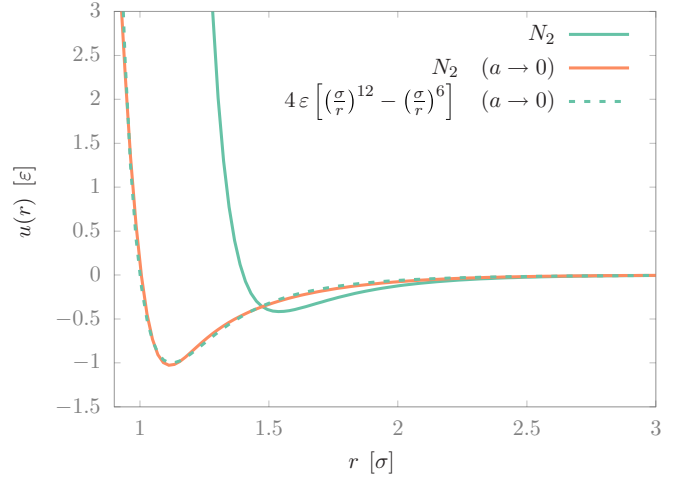


FIG. 1. Comparison between the two-body spherical symmetric potential N_2 produced by the NN, and four times the Lennard-Jones potential to which converges the full diatomic potential when $a \rightarrow 0$. Lennard-Jones diatomic parameters are $a = 2^{1/6}\sigma$. We note that these two potentials do not agree as expected. Three neurons in the hidden layer were used.

III. RESULTS

We test our method using a Lennard-Jones diatomic potential, where all atoms are allowed to interact with

$$u(r) = \epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]. \quad (13)$$

Each molecule is composed of two atoms that are constrained to be part of the diatomic molecule at a distance $a = 2^{1/6}\sigma$, the minimum of the potential. Then, the interaction between two molecules (each composed of two Lennard-Jones atoms) can be calculated with four Lennard-Jones evaluations. Note that if $a \rightarrow 0$ the full interacting potential becomes four times the Lennard-Jones potential, consequently in such case there is no coarse-graining and the approach becomes the one presented in Refs. [31,32]. Using the adaptive algorithms discussed above we construct the virial coefficients as a function of temperature for the non-spherically symmetric potential that we use as data for the construction of the spherically symmetric potential through the NN representation.

We use the same adaptive algorithms to perform an inversion of the second virial data. In Fig. 1 we show the spherically symmetric potential obtained from the inversion procedure of the virial data for the diatomic molecules with $a = 2^{1/6}\sigma$. This potential is compared with the full diatomic potential in the known limit $a \rightarrow 0$ (only in the case when it becomes spherically symmetric). It is interesting to confirm that the two potentials do not agree. Note that the repulsive part of the reduced potential extends to bigger distances than the limit $a \rightarrow 0$. This is expected because it represents in essence a weighted average that includes the fact that the atoms in a molecule at distance a in some configurations can get very close together with other atoms in other molecules. In Fig. 2 we observe a comparison between the second virial data obtained from the Lennard-Jones diatomic potential and the inverted potential. Note that the curves are almost the same

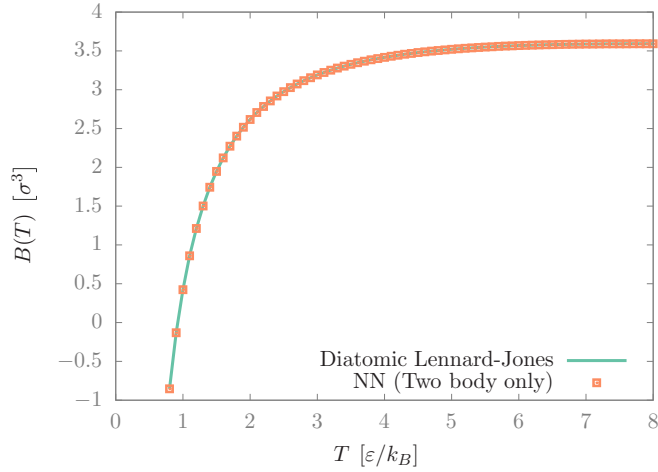


FIG. 2. Comparison between the second virial of the spherical symmetric potential obtained from a Lennard-Jones diatomic potential and the second virial of such potential.

with $S_2 = 10^{-10}\sigma^3$. We then perform an inversion of the third virial data. In Fig. 3 we show the potential obtained in the inversion and in Fig. 4 we observe a comparison between the third virial data obtained from the Lennard-Jones diatomic potential and the inverted spherically symmetric potential. Note that the curves are very close, with a value of $S_2 = 10^{-5}\sigma^6$. At first sight, such an agreement may seem no surprise at all, after all the inversion procedure is constructed to try to ensure so. However, there is no guarantee that a monoatomic potential, that is able to reproduce the same thermodynamics and molecular dynamics, even exists. In this example, we see that such a solution can indeed exist.

Finally, in Fig. 5 we see a comparison between the microscopic potential energy of a fluid of diatomic molecules and a fluid of monoatomic particles interacting with our two-body spherically symmetric potential. Both the calculations were made by a molecular dynamics simulation at fixed temperatures, namely $T = 1.25, 1.5, 2.0 [\epsilon/k_B]$, which were

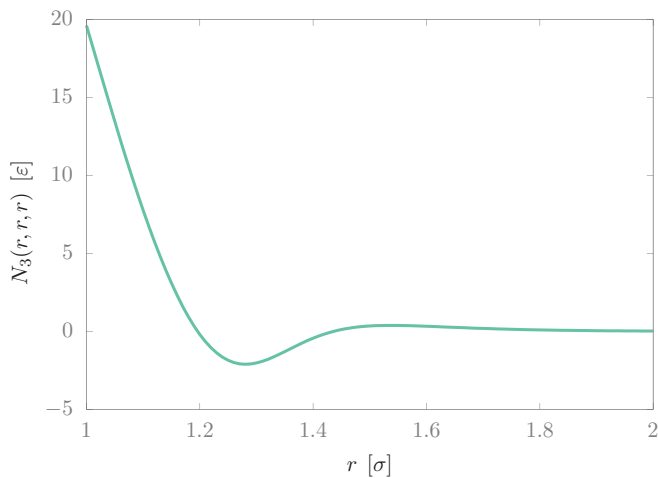


FIG. 3. Three-body spherical symmetric potential N_3 , evaluated at $r = r_{1,2} = r_{1,3} = r_{1,3}$, obtained from a Lennard-Jones diatomic potential. Three neurons in the hidden layer were used.

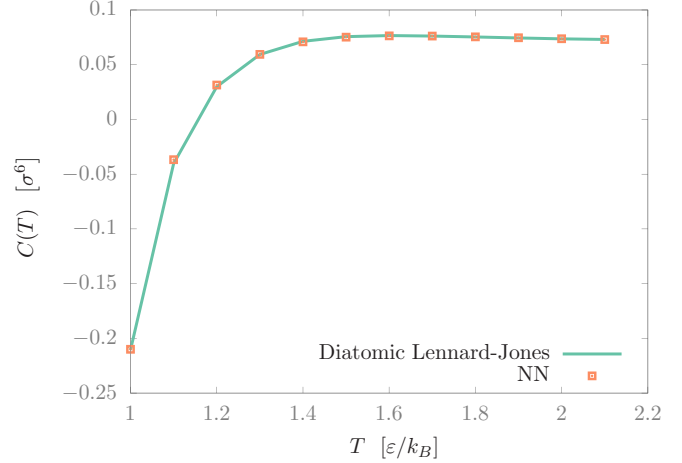


FIG. 4. Comparison between the third virial of the spherical symmetric potential obtained from a Lennard-Jones diatomic potential and the third virial of such potential.

the values of T used to generate the virial data in the training process of the NN. To calculate the internal energy we first got rid of the transient by iterating the system for 100,000 time steps and then we averaged the total energy of the system for an additional 100,000 time steps. We use a time step of $\delta t = 0.003 [\sigma^{-1}\epsilon^{1/2}m^{-1/2}]$ as it was recommended by Johnson *et al.* [37] for Lennard-Jones fluid. In the diatomic case, 1000 molecules were followed with periodic boundary conditions, while in the monoatomic case we use 4000 particles. As a reference, we also present the microscopic potential energy predicted by the virial expansion truncated up to quadratic order in ρ . Note that the correspondence is very good for small ρ and deviation are observed $\rho \rightarrow \sigma^{-3}$ as expected. Moreover, our spherically symmetric potential gives better

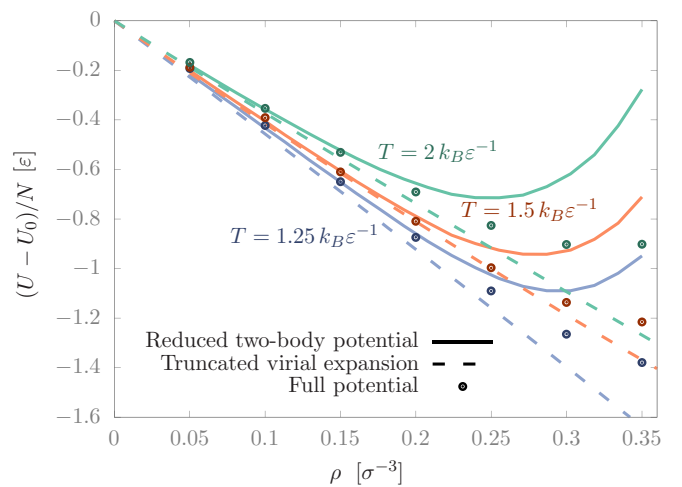


FIG. 5. Comparison between the microscopic potential energy of a fluid of diatomic molecules (dots), the microscopic potential energy of a fluid of monoatomic molecules interacting by our two-body spherically symmetric potential (continuous line), and the internal energy predicted by the virial expansion truncated up to quadratic order in ρ (dashed line), for different temperatures as a function of the number density.

results than the proper virial expansion used for training, and the correct microscopic potential energy always stays between the two results. This can be explained by remembering that the two body potential contributes to all terms in the virial expansion, so that there are effects on the molecular dynamics by the two-body potential beyond the ρ^2 expansion. Although this result can be improved as we add more n -body potentials, it is quite impressive that a spherical potential can reproduce the thermodynamic of a non-spherically symmetric potential, such as the Lennard-Jones diatomic potential used here, and at a reduced computational expense.

IV. CONCLUSIONS AND DISCUSSION

Summing up, we present a systematic reduction of the PES by proposing a simplified version of the potential that preserves the thermodynamics of the system, at least for low density, that can be constructed through successive inversion of the virial coefficients. This is supported by the fact that such inversion is ill-posed, so that there exists more than one potential that reproduces the same virial data, and as such it makes sense to chose a spherical symmetric potential which allows to speed up calculations giving a simpler description of the system. In the future, we plan to use a similar approach to construct a high-density expansion, such as the one used

in Lee-Yang lattice gases [33], to overcome the limitations of our current strategy given by the convergence of the standard virial expansion for larger densities.

In general, we would like to apply our strategy to construct interacting potentials which would allow us to make faster simulations of real liquids or melts of molecules with several degrees of freedom, taking care of preserving the thermodynamics of the system.

Finally, we must note that our approach to reduce the degrees of freedom could be applied not only to the inversion of other thermodynamic quantities, but also to other ill-posed problems, because the existence of many solutions allow us to choose one branch of solutions, in particular, the one that makes the problem simpler.

ACKNOWLEDGMENTS

This work was supported by CONICYT-PCHA/Doctorado Nacional/2016-2116103 (S.C.) and the Fondo Nacional de Investigaciones Científicas y Tecnológicas (FONDECYT, Chile) under Grants No. 1120399 and No. 1130272 (J.R.), No. 1150718 (J.A.V.), and CEDENNA through “Financiamiento Basal para Centros Científicos y Tecnológicos de Excelencia-FB0807” (J.R. and J.A.V.).

-
- [1] D. Frenkel and B. Smit, *Understanding Molecular Dynamics Simulations* (Academic, New York, 2001).
 - [2] B. Smit, P. Hilbers, K. Esselink, L. Rupert, N. Van Os, and A. Schlijper, *Nature* **348**, 624 (1990).
 - [3] R. D. Groot and K. L. Rabone, *Biophys. J.* **81**, 725 (2001).
 - [4] M. Venturoli and B. Smit, *PhysChemComm* **2**, 45 (1999).
 - [5] J. C. Shillcock and R. Lipowsky, *J. Chem. Phys.* **117**, 5048 (2002).
 - [6] W. K. den Otter and W. J. Briels, *J. Chem. Phys.* **118**, 4712 (2003).
 - [7] M. J. Stevens, *J. Chem. Phys.* **121**, 11942 (2004).
 - [8] Laura J. Douglas Frink and A. L. Frischknecht, *Phys. Rev. E* **72**, 041923 (2005).
 - [9] R. Pool and P. G. Bolhuis, *Phys. Chem. Chem. Phys.* **8**, 941 (2006).
 - [10] V. Molinero and E. B. Moore, *J. Phys. Chem. B* **113**, 4008 (2009).
 - [11] A. Reinhardt and J. P. K. Doye, *J. Chem. Phys.* **136**, 054501 (2012).
 - [12] M. Ambler, B. Vorselaars, M. P. Allen, and D. Quigley, *J. Chem. Phys.* **146**, 074701 (2017).
 - [13] H. Noguchi and M. Takasu, *Phys. Rev. E* **64**, 041913 (2001).
 - [14] O. Farago, *J. Chem. Phys.* **119**, 596 (2003).
 - [15] M. Müller, K. Katsov, and M. Schick, *Biophys. J.* **85**, 1611 (2003).
 - [16] G. Brannigan and F. L. H. Brown, *J. Chem. Phys.* **120**, 1059 (2004).
 - [17] T. Murtola, E. Falck, M. Patra, M. Karttunen, and I. Vattulainen, *J. Chem. Phys.* **121**, 9156 (2004).
 - [18] G. Sevink and A. Zvelindovsky, *Macromolecules* **38**, 7502 (2005).
 - [19] I. R. Cooke, K. Kremer, and M. Deserno, *Phys. Rev. E* **72**, 011506 (2005).
 - [20] Z.-J. Wang and D. Frenkel, *J. Chem. Phys.* **122**, 234711 (2005).
 - [21] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein, *J. Phys. Chem. B* **105**, 4464 (2001).
 - [22] S. J. Marrink, A. H. De Vries, and A. E. Mark, *J. Phys. Chem. B* **108**, 750 (2004).
 - [23] S. Izvekov and G. A. Voth, *J. Phys. Chem. B* **109**, 2469 (2005).
 - [24] D. Reith, M. Pütz, and F. Müller-Plathe, *J. Comput. Chem.* **24**, 1624 (2003).
 - [25] S. T. John and G. Csányi, *J. Phys. Chem. B* **121**, 10934 (2017).
 - [26] S. Izvekov and G. A. Voth, *J. Chem. Phys.* **123**, 134105 (2005).
 - [27] A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* **52**, 3730 (1995).
 - [28] M. S. Shell, *J. Chem. Phys.* **129**, 144108 (2008).
 - [29] D. Rosenberger, M. Hanke, and N. F. van der Vegt, *Eur. Phys. J. Special Topics* **225**, 1323 (2016).
 - [30] H. Wang, C. Junghans, and K. Kremer, *Eur. Phys. J. E* **28**, 221 (2009).
 - [31] J. Braga, M. de Almeida, A. Braga, and J. Belchior, *Chem. Phys.* **260**, 347 (2000).
 - [32] R. Sebastião, N. Lemes, L. Virtuoso, and J. Braga, *Chem. Phys. Lett.* **378**, 406 (2003).
 - [33] M. V. Ushcats, L. A. Bulavin, V. M. Sysoev, and S. J. Ushcats, *Phys. Rev. E* **94**, 012143 (2016).
 - [34] K. Fuchs, *Proc. R. Soc. Lond. A* **179**, 408 (1942).
 - [35] E. A. Mason and T. H. Spurling, *The Virial Equation of State*, 1st ed. (Pergamon Press, Oxford/New York, 1969).
 - [36] S. Carrasco, A. Varas, J. Rogan, M. Kiwi, and J. A. Valdivia, *Phys. Rev. B* **94**, 075435 (2016).
 - [37] J. K. Johnson, J. A. Zollweg, and K. E. Gubbins, *Mol. Phys.* **78**, 591 (1993).