Short communication

Thermodynamic consistency in the structure-based integral equation coarse-grained method

M. Dinpajooh, M.G. Guenza[*]

Department of Chemistry and Biochemistry, Institute of Theoretical Science, University of Oregon, Eugene, OR 97403, United States

ABSTRACT

The Integral Equation Coarse-Grained method is an approach that simplifies the representation of a polymer melt into a liquid of coarse-grained chains, considerably speeding up the computation of the melt properties, while reproducing with accuracy structure and thermodynamics of the corresponding atomistic description. In a recent paper [Polymer **111**, 103 (2017)], it was stated that the structure-based Integral Equation Coarse-Grained approach does not give thermodynamic consistency. Here we present new calculations that confirm the validity of this method in predicting consistent pressure and radial distribution functions with atomistic simulations. Other details of the method are also discussed.

© 2017 Elsevier Ltd. All rights reserved.

In the recent paper by Wang and Yang, referred hereafter as Paper I [1], the authors state that the thermodynamic consistency of the atomistic and coarse grained (CG) descriptions of homopolymer melts is not possible when our coarse-graining theory, the Integral Equation Coarse-Graining (IECG) theory, is used. In this paper we clarify the issue of thermodynamic consistency and also address other points that are misinterpretations of our work, as they appear in Paper I. Because it would be impossible to address all the details in the incorrect interpretations of our work as they appear in that paper, we refer the interested reader to our original work, about which we will welcome any questions [2–7]. Many of the criticisms in Paper I were already addressed in a previous paper of ours [8], and will not be addressed again here. In a nutshell, when our approach is used incorrectly and/or outside the limits in which it applies, thermodynamic consistency is lost [8]. However, when the approach is used correctly, thermodynamic consistency holds, as we are going to discuss here.

The original IECG formalism is already described extensively in a number of our publications, so it will be summarized only briefly here. The IECG model is a coarse-graining theory based on the solution of the Ornstein-Zernike integral equation [9], extended to macromolecular liquids [11,12]. In the IECG model each polymer in the melt is described as a chain of CG units, positioned at a distance along the chain comparable or larger than the persistence length [13]. Each polymer is partitioned into a number, n_b, of CG units or blobs, with each CG unit representing a number of monomers, N_b, such that the total number of monomers in a chain is $N = n_b N_b$. In the case where $n_b = 1$, the polymer is described as a point particle interacting with a soft potential, i.e. the soft sphere model. The larger the level of coarse-graining, the more significant the speed up in the Molecular Dynamics (MD) simulations that use the IECG representation and potential. Note that Gaussian statistics, which is used in the analytical solution of the IECG theory, is satisfied when N_b is roughly greater than 30 for a melt of polyethylene chains, while different numbers fulfill this condition when polymers of different chemical nature are studied. Also, the direct correlation function, and so the potential, between blobs involves a nonzero tail at relatively large distances, which is more pronounced as the level of coarse-graining increases [3,4]. Structural properties, i.e. the pair distribution function $h(r) = g(r) - 1$, and thermodynamic properties, i.e. compressibility, pressure, internal energy, entropy, Gibbs excess free energy, and Helmholtz excess free energy, were calculated in our previous works for polyethylene melts at different densities and degrees of polymerization [2–7]. Theoretical predictions and results from coarse-grained IECG simulations were compared and tested against united atom (UA) simulations. Structural and thermodynamic properties are consistent, with the exception of the entropy and the internal energy, which depend on the number of blobs in which a chain is partitioned.

We start by addressing with an example the question if it is

---

* Corresponding author.
   E-mail address: mguenza@uoregon.edu (M.G. Guenza).

possible for the CG simulations, which use the structurally-consistent IECG potential, to predict thermodynamic properties consistent with the corresponding atomistic simulations. Reported in Fig. 1 are the results of the pressure calculated in four different MD simulations, performed with different levels of granularity: i) a United Atom simulation (reported in red in the figure), ii) a simulation where each polymer in the melt is represented as a soft sphere ($n_b = 1$, reported in blue), iii) a simulation in which each polymer is represented as a chain of 4 blobs ($n_b = 4$, reported in orange), and iv) one simulation in which each polymer is represented as a chain of 6 blobs ($n_b = 6$, reported in cyan). Each simulation represents a melt of polyethylene chains, with a degree of polymerization $N = 192$, at the temperature of 509 K and at a monomer density of 0.733 gr/cm^3. All the simulations are performed in the canonical ensemble with the Nose′-Hoover thermostat, using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software program [14].

The UA simulation results are based on the analysis of 8 ns trajectories available from our previous work [3,6]. More details about the UA simulations are reported there. The CG simulations were performed using the IECG potential. The timestep for all was set to 10 fs and the temperature in the Nose′-Hoover thermostat was relaxed in a timespan of 200 fs. The CG particles were allowed to equilibrate for 1 ns before production runs of $6 − 12$ ns were performed to obtain static properties and thermodynamic averages. The IECG potential depends on one non-trivial parameter, the effective monomer-monomer direct correlation function in the limit of zero wavevector k, $C_0 = C^{mm}(k \to 0) = 4\pi \int_0^\infty r^2 C^{mm}(r) dr$, which is directly related to the compressibility of the system. The specific value of C_0 can be determined by different procedures, as described below. In the simulations presented here the value of $C_0 = -10.02$ Å^3. Another important, but trivial, parameter is the polymer squared radius of gyration, which is assumed to be $R_g^2 = 541.0$ Å^2. This value of the radius of gyration was selected from data of Neutron Spin Echo experiments, performed by Richter and coworkers at the given thermodynamic conditions of

temperature and density [15]. This value is consistent with a straightforward theoretical estimate given by $R_g^2 = Nl^2/6$, where l is the effective segment length between the center-of-mass of two monomers (4.37 Å for polyethylene). When the theoretical value is used, this results in a R_g^2 value of 611.1 Å^2 for $N = 192$. Simulations performed with this value of R_g^2 and the chains coarse-grained, for example, with six blobs, $n_b = 6$, gives a pressure of 356.15 ± 0.03 atm, which is in excellent agreement with the results obtained using $R_g^2 = 541.0$ Å^2, i.e. a pressure of 356.12 ± 0.03 atm, as reported in Fig. 1.

The cutoff distance for the CG simulations was determined over the range of $r > 0$ from the location of the second extremum of effective intermolecular potential, $V^{bb}(r)$, obtained from the Hypernetted Chain closure (HNC), noting that a small attractive part exists between the first and second extrema [4]. The cutoff values of 269.48 Å, 102.26 Å, and 77.07 Å were used for n_b of one, four, and six, respectively. The simulations with $n_b = 1$ were performed in a box size of 547.8 Å consisting of 27000 blobs, while the simulations with n_b of 4 and 6 were performed in a box size of 273.9 Å, each consisting of 13500 and 20250 blobs, respectively. For n_b of 4 and 6, the effective bond potential between adjacent blobs was taken to be $V_{bond}(r) = 3n_b k_B Tr^2/(8R_g^2) + V^{bb}(r) + k_B T \ln[1 + h^{bb}(r)]$, which enforces the appropriate distributions between adjacent blobs, where k_B is the Boltzmann constant and T is the temperature. The angle potential was calculated using $V_{angle}(\theta) = -k_B T \ln[P(\theta)/\sin(\theta)]$, where Eq. (33) of McCarty et al. [6] was used to obtain $P(\theta)$. All $1 − 2$ and $1 − 3$ pairs of bonded blobs were excluded from the non-bonded interactions.

The top panel of Fig. 1 shows the evolution of the instantaneous pressure, P_i at step i, while the black line represents its instantaneous average: for example at step j, $\langle P_j \rangle = \frac{1}{N_j} \sum_{i=1}^{N_j} P_i$ as N_j is the number of steps up to step j. The bottom panel of Fig. 1 shows in red the probability distribution of the pressure, as obtained from the UA simulation, with the average pressure in black, and the
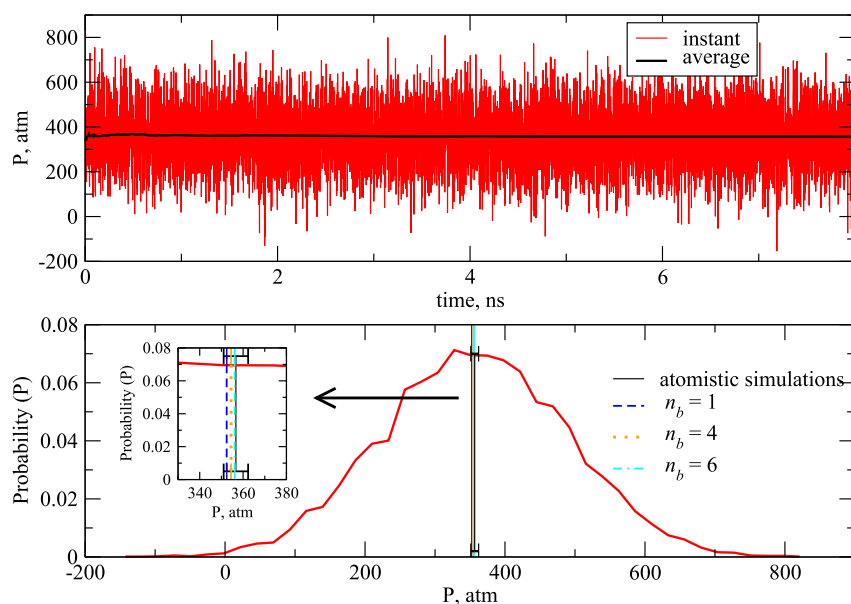

Fig. 1. Top panel: evolution of the instantaneous pressure over the production run of an atomistic simulation of polyethylene with $N = 192$, at 509 K and density 0.733 gr/cm^3 (red curve). Average pressure for the same simulation (black curve). Bottom panel: pressure distribution for the atomistic simulation (red curve). The black line shows the average pressure for atomistic simulations with the simulation error bars (99% confidence interval) obtained from block averages. The dash blue, dot orange, and dash-dot cyan lines show the average pressure for coarse-grained (CG) simulations with one, four, and six CG units per chain, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

simulation error bar (99% confidence interval), where the Standard Error of the Mean (SEM) was obtained from 12 block averages of the simulation run and multiplied by 3.106 to obtain the 99% confidence intervals. The inset shows the average pressure measured in the MD simulations of the CG polyethylene melt, with the CG polymers interacting through the IECG potential, when the chains are represented by one (dash blue line), four (dot orange line), and six (dash-dot cyan line) CG units. The error bar in the CG representations is of the order of the line width. *All the CG pressures fall within the error of the UA simulation.*

With regard to the issue of thermodynamic consistency, our key point of disagreement with Paper I is the following: the authors of Paper I argue that thermodynamic consistency can be claimed only when $A = B$, given that A and B are the thermodynamic properties of the atomistic and CG systems, respectively. They also ruled out the use of molecular simulation results as a proof of thermodynamic consistency of a coarse-grained model. We disagree with them because thermodynamic consistency does not necessarily mean a mathematical identity, i.e., $A = B$. In fact, we believe thermodynamic consistency is achieved when the simulations from different levels of coarse-graining yield results that are within the statistical uncertainties of the atomistic simulation itself, as shown in Fig. 1.

We also confirm that the mathematical identity between atomistic and coarse-grained thermodynamic properties is in fact obtained by the analytical solution of the IECG theory under certain conditions. In this respect, it is important to notice that integral equation theories, as any theory that addresses the physics of complex molecular fluids, are not exact and that approximations are always implicit in the analytical solution of coarse-grained models. Specifically in our approach, the mathematical identity of coarse-grained and atomistic properties is obtained from the solution of the Ornstein-Zernike equation in the conditions where a mean-field equation between the direct correlation function, $C(r)$, and the potential, $V(r)$, applies, i.e. $C(r) = -\beta V(r)$ with $\beta = (k_B T)^{-1}$ [9]. This is the so-called "Mean Spherical Approximation" (MSA) closure, which can be derived formally from the HNC closure, i.e. from the equation used in the numerical calculation of the potential, by assuming that the pair distribution function is equal to one. When the polymer is coarse-grained, the potential is bound, and is represented by a soft, long-ranged, and a smoothly-varying function: for example, for a liquid of soft spheres ($n_b = 1$) of chain density ρ_c, let us define the range of the potential as σ. Then we have $\rho_c \sigma^3 \gg 1$, while the average interparticle distance is $a = \rho_c^{-1/3}$, from which follows that $\sigma \gg a$. In these conditions, each CG site is interacting with a very large number of surrounding CG particles and a mean-field description applies. The excess free energy of the system can be well approximated by the mean-field equation $F_{ex}[\rho_c(\mathbf{r})] \approx 1/2 \iint d\mathbf{r}d\mathbf{r}' V^{cc}(|\mathbf{r} - \mathbf{r}'|)\rho_c(\mathbf{r})\rho_c(\mathbf{r}')$, where $V^{cc}(r)$ is the potential and $\rho_c(\mathbf{r})$ is the position-dependent density profile. Because the potential is a slowly decaying function and the liquid is homogeneous, in the range of the potential σ the density can be considered constant. Given its definition, this implies that the direct correlation function becomes independent of the density

$$C^{cc}(|\mathbf{r} - \mathbf{r}'|; \rho_c) = - \lim_{\rho_c(\mathbf{r}) \to \rho_c} \frac{\beta \partial^2 F_{ex}[\rho_c(\mathbf{r})]}{\partial \rho_c(\mathbf{r}) \partial \rho_c(\mathbf{r}')} \tag{1}$$

recovering the MSA, $C^{cc}(r) = -\beta V^{cc}(r)$ [10]. It follows that the potential, solved in the MSA, and the equation of state derived from it, are not in contradiction with the structural consistency of the pair distribution function.

Furthermore, it is evident that the use of the MSA is well justified when the potential is a long-ranged, slowly-varying function,

for which the mean-field approximation of the excess free energy becomes correct. Long-ranged, slowly-varying potentials are obtained for highly coarse-grained models and for almost incompressible liquids where the density of CG particles is high. In these conditions, the IECG potential and thermodynamic properties are solved analytically, demonstrating the consistency in structure, compressibility, pressure, and excess free energy (see for example Eq. (3)). On the other hand, the analytical solutions of the internal energy and of the entropy show a dependence on the number of CG sites that model the polymer. This dependence is expected, given that the process of coarse-graining averages out a number of local degrees of freedom, modifying the entropy of the system.

The analytical solution of the pressure is calculated using the virial expression

$$\frac{P}{\rho_c k_B T} = 1 - \frac{2\pi\rho_c}{3k_B T} \int_0^\infty r^3 g^{cc}(r) \frac{\partial V^{cc}}{\partial r} dr, \tag{2}$$

where P is the pressure and $\rho_c N = \rho_m$, where ρ_m is the monomer density. From direct inspection of the integral, one can see that the tail of the pair distribution function and the tail of the force are dominating the integral due to the r^3 factor. Because the potential of a coarse-grained description has a range, σ, that is much larger than the distance of the first solvation shell, given by the interparticle separation a, it is a valid approximation to assume in the calculation of the virial pressure that $g(r) \approx 1$ for liquids of coarse-grained polymers. Applying this approximation together with the MSA closure leads to the equation of state

$$\frac{P}{\rho_c k_B T} = 1 - \frac{\rho_m N C_0}{2}. \tag{3}$$

Eq. (3) does not depend on the number of blobs a chain is partitioned, n_b, or on the number of monomers in one blob, N_b; thus, the pressure calculated from Eq. (3) is independent of the level of granularity adopted in the coarse-grained model.

Next, we check the validity of Eq. (3) by straightforward numerical calculations. We show here that the relative error introduced by the assumption $g^{cc} = 1$ ($n_b = 1$) is indeed about 0.2% when the pressure obtained by this route is compared to the one that does not use the approximation. The calculation is performed for the sample studied in this paper ($N = 192$) in the soft sphere model, $n_b = 1$. Integration by parts of the second term on the right hand side in Eq. (2) gives

$$\frac{2\pi\rho_c}{k_B T} \int_0^\infty r^2 \left(g^{cc}(r) + r \frac{\partial g^{cc}(r)}{3\partial r} \right) V^{cc}(r) dr, \tag{4}$$

which, using the HNC closure reduces to

$$\frac{\rho_m N}{2} \frac{4\pi}{N^2} \int_0^\infty r^2 \left(g^{cc}(r) + r \frac{\partial g^{cc}(r)}{3\partial r} \right) \beta V_{HNC}^{cc}(r) dr = -\frac{\rho_m N \Lambda_{0HNC}}{2}, \tag{5}$$

where, $\beta V_{HNC}^{cc}(r) = \ln(g^{cc}(r)) - h^{cc}(r) + C^{cc}(r)$. In the MSA, one obtains

$$-\frac{\rho_m N}{2}\frac{4\pi}{N^2}\int_0^\infty r^2\left(g^{cc}(r)+r\frac{\partial g^{cc}(r)}{3\partial r}\right)C^{cc}(r)dr = -\frac{\rho_m N \Lambda_{0MSA}}{2},$$

(6)

where, $C^{cc}(r)$ is the blob direct correlation function. Both Eq. (5) and Eq. (6) reduce to the following when $g^{cc}=1$ is assumed

$$-\frac{\rho_m N}{2}\frac{4\pi}{N^2}\int_0^\infty r^2 C^{cc}(r)dr = -\frac{\rho_m N C_0}{2},$$

(7)

where $C_0 = N^{-2}C^{cc}(k\to 0)$. Here we numerically compare the values of Λ_{0HNC}, Λ_{0MSA}, and C_0 in Eqs. (5)–(7), respectively, for the soft sphere model ($n_b = 1$), noting that $C^{cc}(r)$ is more long-ranged than $g^{cc}(r)$. We obtain values of $\Lambda_{0MSA} = -9.995$, $\Lambda_{0HNC} = -10.003$, and $C_0 = -10.016$ Å³, which result in the pressure values of 355.78, 356.08, and 356.52 atm, respectively. Therefore, the relative error in predicting the pressure is about 0.2% and 0.1% for MSA and HNC, respectively, when approximation $g^{cc} = 1$ is used for the soft sphere model in this work.

The thermodynamic consistency shown in Fig. 1 is not limited to the example that we report above, but is a general property of the IECG approach in the conditions in which the present theory applies. As we discussed earlier, those conditions are, roughly speaking, melt density, and temperatures comparable to the ones reported in the example of Fig. 1. If our theory is erroneously used in conditions in which the IECG approach does not apply, e.g. for polymers in solution or low densities, the IECG approach in its present version cannot ensure thermodynamic consistency [8].

Some other points worth reporting are:

i) In this work, the structural consistency is also shown for the blob-blob radial distribution function, when the polymer is represented by one or four blobs (see Fig. 2). As can be seen, the difference between the atomistic and CG simulations

results lies within the statistical uncertainties of the atomistic simulations. In addition, both are in excellent agreement with the theoretical prediction of the IECG. This indicates that in the IECG method, the closure, together with the Ornstein-Zernike equation from the liquid state theory, implicitly capture with accuracy many-body effects. Without correctly capturing the many-body contributions to the potential, the excellent agreement that is observed would not be possible.

ii) The IECG approach is structurally and thermodynamically consistent in pressure and excess free energy. This consistency is important because a CG model, to be useful, needs not only to speed up the computational time, but also to be consistent in its predictions of structural and thermodynamic quantities: fast CG simulations that predict incorrect thermodynamics are of limited use.

iii) The IECG formalism is analytical in the sense that thermodynamic and molecular parameters explicitly enter the potential. In this way the IECG potential is *transferable*, as it applies to other sets of conditions, within the range of parameters discussed above, and to polymers with different chemical structures, while it predicts structural and thermodynamic properties consistent with the related atomistic simulations [2–6].

iv) While the IECG potential depends on one non-trivial parameter, C_0, which is directly related to the compressibility of the system, this should not be a source of concern. Note that the potentials commonly used in atomistic simulations depend on more than one non-trivial parameter (for example, the Lennard-Jones potential depends on two parameters). The value of C_0 can be obtained in a number of different ways that give numerically consistent values. For example, the value of C_0 used in this work (-10.02 Å³) may be obtained from the derived equation of state, Eq. (3). To derive C_0 from the equation of state, the pressure has to be known for a small number of samples, and not necessarily for the one for which the coarse-grained simulation is
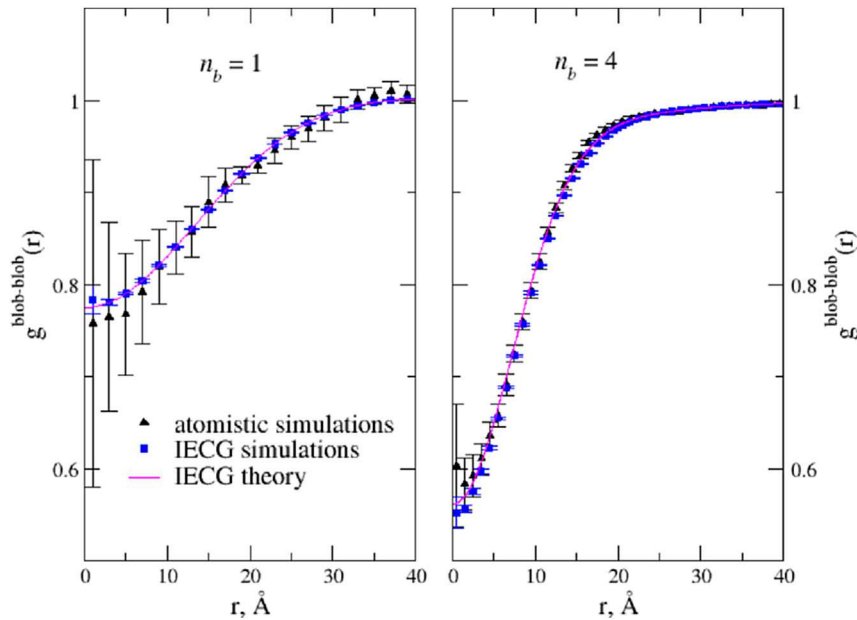

Fig. 2. Theoretical and simulated radial distribution functions when the polymer is represented by one (left panel) and four (right panel) coarse-grained sites. Atomistic simulations (black triangle) are compared with coarse-grained simulations (blue square) and with the analytical function (magenta line). The analytical expression and the data from the coarse-grained simulation are both within the error of the atomistic simulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

performed. In this way, it is not necessary to first perform atomistic simulations of the sample of interest to derive the coarse-grained potential, which sets this method apart from all the other coarse-grained methods. We also note that the C_0 value obtained by the above route from the equation of state is in close agreement with the one reported in Figure 5 in Ref. [6].

iv) Often potentials that are optimized numerically to reproduce the structure of a liquid are not able to reproduce the pressure or the compressibility of the same. This is a well-known result, consistent with the predictions of Statistical Mechanics [9]. These results, however, do not apply to our approach, where many-body interactions are accounted for by the Ornstein-Zernike equation. Because coarse-graining reduces the number of degrees of freedom, thermodynamic consistency in the IECG approach applies only to quantities that do not depend on the number of degrees of freedom. Thus, the IECG reproduces with accuracy pressure and free energy, while entropy and internal energy are modified [3,4,6]. The fact that entropy and internal energy are not conserved between the atomistic and the coarse-grained multiblob description is, in fact, not a limitation of our method but a result supported by the analytical solution of the IECG model. The analytical solution of the IECG approach shows that both entropy and internal energy formally depend on the number of blobs in which the chain is partitioned, n_b. It follows that those properties should not be conserved during coarse-graining, as can be intuitively understood by the fact that the number of degrees of freedom associated with the CG model changes with the granularity of the model.

v) Paper I often cites, in the same sentence, words that are taken from two different papers of ours: the first is the paper by Mc Carty et al. [5] (reported as Reference 4 in Paper I), and the second is the paper by Clark et al. [4] (reported as Reference 5 in Paper I). One should keep in mind that the two papers refer to two different intramolecular polymer chain models. The first paper by McCarty et al. [5], as well as the paper by Yatsenko et al. [7] also cited in Paper I, is based on the thread model [12] for which the structure and the thermodynamic properties are solved *analytically*: for this model the thermodynamic consistency is satisfied in the corresponding limits. The second paper by Clark et al. [4] is based on the Freely Jointed (FJ) Chain representation of the polymer chain, which has been solved analytically in the two limit of large intermolecular distances between pairs of CG units and small intermolecular distances. The numerical solution of the FJ IECG model is valid at any distance, and its results are reported in our papers, and in Figs. 1 and 2 above.

Sentences and methods from the paper by McCarty et al. [5], do not always apply to the theory, the system, and the results reported in the paper by Clark et al. [4], and vice versa. For example, in Ref. 4 of Paper I [5], thermodynamic consistency is only obtained for the *interchain* contribution to the internal energy and the *interchain* contribution to the virial pressure, because *each polymer is coarse-grained to a point particle and, in that case, no intramolecular contributions are present*. It is always true that both intra and intermolecular contributions need to be included in the energetics and in the pressure when the polymer is coarse-grained as a chain of blobs, otherwise those quantities would not be consistent in variable levels of representation [8].

vi) A careful and useful analysis of the contributions to the thermodynamic properties in our theory is reported in Figure 7 of Clark et al. [3], where the figure shows how different terms contribute to the final pressure of the liquid. The discussion following Figure 7 of Clark et al. [3] shows that the IECG theory is most accurate when the liquid is dense and the chains are long, while N_b is larger than the polymer persistence length. This is, in fact, the type of system for which coarse-graining models become necessary when performing a simulation because atomistic simulations are slow and can require excessive computational resources. In the IECG approach, all contributions to the thermodynamics are included. Pressure, for example, contains contributions from kinetic, intramolecular, and intermolecular terms, as explained in the discussion following Figure 7 of Clark et al. [3].

In this paper, thermodynamic consistency of the IECG model is validated for melts of long polymer chains, for which it is possible to perform MD simulations with variable levels of coarse-graining, while conserving the pressure, the compressibility, and the excess free energy. This is a clear advantage of the IECG method, with respect to other approaches.

References

[1] Q. Wang, D. Yang, Polymer 111 (2017) 103.
[2] A.J. Clark, M.G. Guenza, J. Chem. Phys. 132 (2010) 044902.
[3] A.J. Clark, J. McCarty, M.G. Guenza, J. Chem. Phys. 139 (2013) 124906.
[4] A.J. Clark, J. McCarty, I.Y. Lyubimov, M.G. Guenza, Phys. Rev. Lett. 109 (2012) 168301.
[5] J. McCarty, A.J. Clark, I.Y. Lyubimov, M.G. Guenza, Macromolecules 45 (2012) 8482.
[6] J. McCarty, A.J. Clark, J. Copperman, M.G. Guenza, J. Chem. Phys. 140 (2014) 204913.
[7] G. Yatsenko, E.J. Sambriski, M.A. Nemirovskaya, M. Guenza, Phys. Rev. Lett. 93 (2004) 257803.
[8] A.J. Clark, J. McCarty, M.G. Guenza, J. Chem. Phys. 143 (2015) 067101.
[9] J.-P. Hansen, I.R. McDonald, Theory of Simple Liquids, Academic Press, London, 1991.
[10] R. Evans, Adv. Phys. 28 (1979) 143.
[11] D. Chandler, J. Chem. Phys. 59 (1973) 2742.
[12] K.S. Schweizer, J.G. Curro, Adv. Chem. Phys. 98 (1997) 1.
[13] P.J. Flory, Principles of Polymer Chemistry, Cornell University Press, Ithaca, United States, 1953.
[14] S. Plimpton, J. Comp. Phys. 117 (1995) 1.
[15] M. Zamponi, A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, P. Falus, B. Farago, M.G. Guenza, J. Phys. Chem. B 112 (2008) 16220.