Evaluating virial coefficients for multicomponent mixtures: hard sphere mixtures and flexible chains

C. VEGA*

Departamento de Química Física, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain

A new algorithm to compute the virial coefficients of multicomponent mixtures is proposed. The number of graphs that must be evaluated increases dramatically in a multicomponent mixture so that it becomes difficult to enumerate and compute all possible graphs. However, once all of them are known and evaluated, the virial coefficient of the mixture can be evaluated for any composition. If one is interested in the virial coefficient of a mixture of a certain composition, then a simpler approach can be followed. Starting from the graphs of a pure fluid, we assign a random chemical identity to each of the molecules of the graph. The probability of assigning a given chemical identity is taken from the composition of the mixture. In this way composition is treated as a random variable within the Monte Carlo procedure which determines the virial coefficient. The algorithm is checked by comparison with the virial coefficients of binary hard spheres mixtures which are well known. Good agreement is found. The procedure is then extended to multicomponent mixtures of hard spheres. Finally the procedure is applied to the determination of the virial coefficients of a flexible molecule. For flexible molecules the possible configurations of the molecules are treated as different components of the mixture. In this way we present what appears to be the first determination of the third and fourth virial coefficients of polymers in the continuum.

1. Introduction

The virial expansion of a pure fluid provides an expansion of the compressibility factor in powers of the number density of the fluid. From a molecular point of view, virial coefficients can be obtained by evaluating integrals over clusters of molecules [1–3]. These integrals are usually represented by graphs and have been well known for a long time. For instance the fourth virial coefficient can be obtained by evaluating several integrals involving clusters of four molecules. Virial coefficients up to the fourth or fifth of pure fluids are now well known for a number of models: e.g. hard spheres [4], hard convex bodies [5–13], models of linear rigid tangent hard spheres [14] and for systems interacting through attractive forces as in the case of the Lennard-Jones [15, 16] system. In the case of mixtures, most of the work has focused on binary mixtures. Probably the binary mixture of hard spheres has been the system for which most work has been devoted to the determination of virial coefficients of mixtures. Virial coefficients of binary mixtures of hard spheres were determined long ago [17], for moderate size ratios and recently for larger size ratios [18–

24]. The interest in the study of binary mixtures of hard spheres for large size ratios can be attributed to the possibility, suggested by Hansen and Biben, of a fluid–fluid phase separation for this system [25]. After one decade of controversy it now seems clear that there is no fluid–fluid separation within the phase diagram of mixtures of additive hard spheres [26–28].

The motivation for this work has two different origins. Let us describe briefly each of them. The first arises from studies of the virial coefficients for binary mixtures of hard spheres. The fifth virial coefficient of hard spheres has recently been determined by Giaquinta and co-workers [19, 20] and Enciso *et al.* [22]. The results of these two groups were not in agreement. The origin of the problem was an incorrect graph [29, 30]. The correct graphs for the fifth virial coefficient of a binary mixture were those given by Enciso *et al.* [22]. Once the correct graphs were used then the results of Wheatley *et al.* [21] showed good agreement with those of Enciso *et al.* [22]. The reason for mentioning this point is just to illustrate that the writing of the graphs of the virial coefficient of high order, as for instance the fifth, or the sixth, may become tricky and mistakes are easily made. Imagine now, that we wish to compute the fifth virial coefficient of a 10-component mixture. The

---

*e-mail: cvega@eucmos.sim.ucm.es

number of possible graphs increases dramatically and since the problem becomes quite complex one should watch carefully for the possibility of a mistake. Even if one finally wrote all the graphs, one would then like to have a different procedure to evaluate the virial coefficient so that the algorithm can be checked. This is the first motivation for this work: to provide a simple algorithm to compute the virial coefficients of a multicomponent mixture. This may be useful per se, or as a cross-check of the determination of the virial coefficients by rigorous evaluation of all possible graphs.

The second motivation for this work arises from our studies of flexible molecules. Flexible molecules adopt a number of configurations or shapes. If one uses the rotational isomeric state (RIS) approximation [31] (where each torsional degree of freedom is discretized into a few torsional states) then the flexible molecule is indeed a multicomponent mixture. Certainly this is a special mixture because composition is not a free variable since all configurations of the molecule must have the same chemical potential. To illustrate the point let us mention the case of n-alkanes. Within the RIS approximation three states are associated with each torsional degree of freedom [32] which are usually denoted as trans (t), gauche+ (g^+) and gauche− (g^-). The number of configurations of n-hexadecane within the RIS approximation is about 3^{13}. In this case the number of components is so huge that one realizes that the problem of evaluating the virial coefficients of n-hexadecane by exhaustive enumeration and evaluation of all graphs is out of the question. If one analyses the literature it can be concluded that evaluation of virial coefficients of flexible molecules of moderate length (i.e. with a number of monomer units higher than 20) have been performed only for the second virial coefficient [33–37], and quite recently for the third virial coefficient of a lattice polymer [38]. No previous determination of the third and fourth virial coefficients of a flexible model in a continuum have been determined. This is the second motivation for this work: to provide a procedure to determine high virial coefficients of polymers. We shall focus on hard bodies since the determination of high virial coefficients (fourth, fifth) of polymers presenting attractive forces is beyond the limits of current modern computers.

One can say that in this work we provide an algorithm to compute virial coefficients of multicomponent mixtures. The number of components can be small (2), moderate (10) or huge (as in the case of a flexible molecule). The reason to join, in the same contribution, mixtures of hard spheres and hard polymers is to illustrate how the procedure applies equally well to quite different systems.

2. Virial coefficients of a multicomponent mixture

The pressure of a homogeneous isotropic mixture can be given in terms of powers of the density by the following expression:

$$Z = \frac{p}{\rho kT}$$
$$= 1 + B_2(\mathbf{x}, T)\rho + B_3(\mathbf{x}, T)\rho^2 + B_4(\mathbf{x}, T)\rho^3 + \cdots, \quad (1)$$

where ρ is the number density (number of molecules per unit volume) of the system and the coefficients B_2, B_3 and B_4 are the second, third and fourth virial coefficients of the mixture respectively. These virial coefficients are a function of the temperature T and of the composition of the system, which is represented by $\mathbf{x}$. In this work we shall deal mainly with hard bodies for which the compressibility factor does not depend on temperature so that virial coefficients are not temperature dependent.

The link between the macroscopic expression given by equation (1) and the microscopic world was made in the 1930s when it was shown that the virial coefficients could be given in terms of certain integrals involving the potential energy between molecules. These integrals are commonly represented by graphs [2, 3]. For a pure fluid (i.e. when the number of components is one) the graphs which yield the virial coefficients are well known. In particular in figure 1 the graphs required to compute the fourth virial coefficient of a pure fluid (we shall assume pairwise potentials throughout this work) are shown. Solid lines represent Mayer functions $f = \exp(-\beta u) - 1$, where $\beta = 1/(kT)$ and u is the pair potential between pairs of molecules. Black circles represent integration with respect to the coordinates of that molecule. Each graph represents the value of a certain integral. The value of the integral is obtained by assigning a number to each black circle of the graph and performing integration with respect to all the variables. For instance the first graph in figure 1 represents the value of the following integral:

$$\int f(1,2)f(2,3)f(3,4)f(4,1)d1d2d3d4, \quad (2)$$
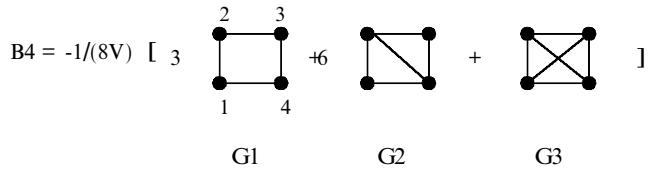
Figure 1. Graphs appearing in the determination of B_4 for a pure fluid with no internal degrees of freedom. Associated with each circle we have the function 1. The diagrams are labelled so that the value of the diagram is just the value of the corresponding integral, without need of introducing the symmetry number of the diagram. The three graphs are denoted as G1, G2 and G3.

$$B4 = -1/(8V) \quad [\quad 3 \quad \text{G1} \quad +6 \quad \text{G2} \quad + \quad \text{G3} \quad]$$
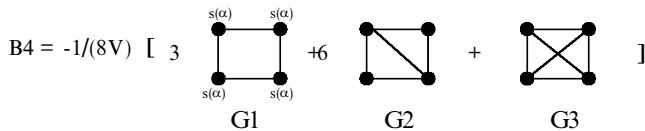
Figure 2. Graphs appearing in the determination of B_4 of a pure fluid with internal degrees of freedom. Associated with each black circle is the internal distribution function $s(\alpha)$. The three graphs are denoted as G1, G2 and G3.

where $d1$ represent the set of coordinates required to define the position and orientation of molecule 1. It should be understood that all the diagrams presented in this work correspond to labelled diagrams (i.e. no symmetry number must be associated to the graph since the factors 3, 6 and 1 appearing in figure 1 already take this into account). The labelling used in this work is illustrated in the first graph of figure 1. Labels 1 and 2 will be attached to the bottom and top left black circles respectively, and labels 3 and 4 will be attached to the top and bottom black circles of the right side (of course the integral does not depend on the arbitrary way of labelling since the integration variables are dummy variables). We shall use this labelling in all the diagrams of this work.

Let us now go from the pure fluid to a more general case. Let us imagine that each molecule of the system may exist in different states (we shall clarify the meaning of the word states later on). Then the fourth virial coefficient of a system where each molecule can exist in different states is given by the graphs shown in figure 2. Figure 2 is identical to figure 1 except that the function $s(\alpha)$ is assigned to each black circle instead of the factor 1 of figure 1.

The function $s(\alpha)$ associated to each black circle will be known as the internal distribution function. The internal distribution function $s(\alpha)$ is the function defined by the equation (for homogeneous isotropic fluids):

$$\rho_\alpha = \rho s(\alpha), \tag{3}$$

where ρ_α is the number density of molecules in the state denoted as α. The function ρ_α is usually denoted as the singlet correlation function. If the number of internal states of the molecule is continuous then $\rho(\alpha)\,d\alpha$ gives the number density of molecules between α and $\alpha + d\alpha$. Let us mention that the term 'state of the molecule' can be interpreted in a number of ways. One may think of it as corresponding to an internal state of the molecule, to a certain configuration in the case of a flexible molecule, or even to a chemical identity in the case of a mixture. From the definition of equation (3) it follows that for the particular case of a rigid pure fluid the function $s(\alpha)$ is

identically one. From the definition given by equation (3) it can be seen that the function $s(\alpha)$ represents the probability of the molecule being in the state α. Therefore the function $s(\alpha)$ is normalized to one and it can be considered as a probability distribution function. When the number of states of the molecule is discrete we have:

$$\sum_i s(\alpha_i) = 1. \tag{4}$$

If the number of states of the molecule is continuous rather than discrete then:

$$\int s(\alpha)\,d\alpha = 1. \tag{5}$$

Once the function $s(\alpha)$ has been defined then the first graph of figure 2 represents the integral:

$$\int f(1,2)f(2,3)f(3,4)f(4,1)s(\alpha_1)s(\alpha_2)$$
$$\times s(\alpha_3)s(\alpha_4)d1\,d2\,d3\,d4\,d\alpha_1\,d\alpha_2\,d\alpha_3\,d\alpha_4. \tag{6}$$

Let us now be more precise about the meaning of the function $s(\alpha)$. For a multicomponent mixture we can assume that each component corresponds to a different state. Therefore the number density of component i is given by:

$$\rho_i = \rho x_i. \tag{7}$$

By comparing equation (7) to equation (3) it can be shown that the internal state distribution function is just the set of molar fractions of the components of the mixture. In the case of a binary mixture of two components A and B, then the internal state distribution function is given by the molar fractions x_A and x_B respectively. The implementation of the algorithm to compute the fourth virial coefficient of a binary mixture from the diagrams of figure 2 is particularly simple and is given by the following steps.

(a) Choose a composition for which one wishes to evaluate the value of B_4.
(b) Randomly select the chemical identity (or state) of molecule 1 from the distribution $s(\alpha)$. For a binary mixture with $x_A = 0.8$ and $x_B = 0.2$ this means that one generates the chemical identity of molecule 1 with 80% probability of being a molecule of type A and 20% of being a molecule of type B. Once the identity of molecule 1 has been chosen, choose the chemical identity of molecule 2, with a probability of 80% of being a molecule of type A and 20% of being a molecule of type B. Repeat this procedure, with molecules 3 and 4.
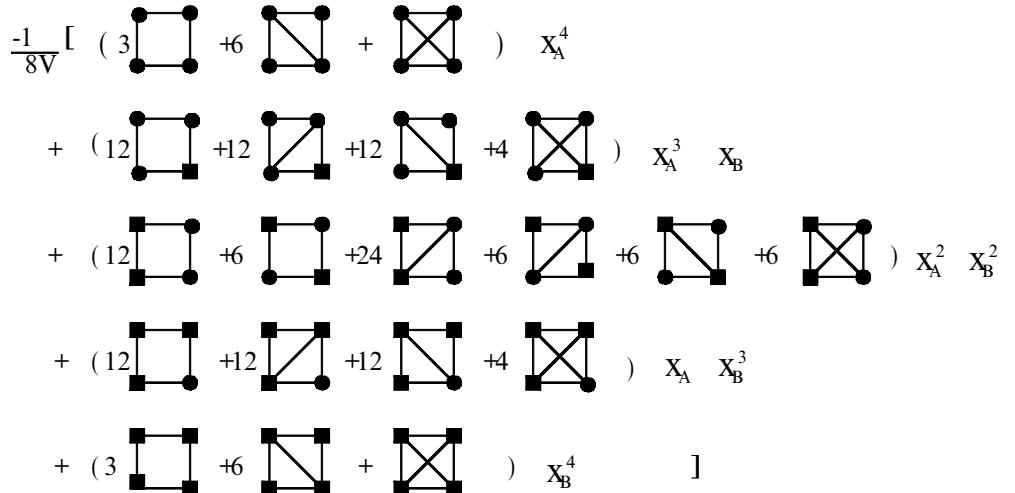
Figure 3. Graphs appearing in the determination of B_4 of a binary mixture in which the two components have no internal degree of freedom. Black circles correspond to component A and black squares to component B. Associated with each symbol (circle or square) we have the function 1.
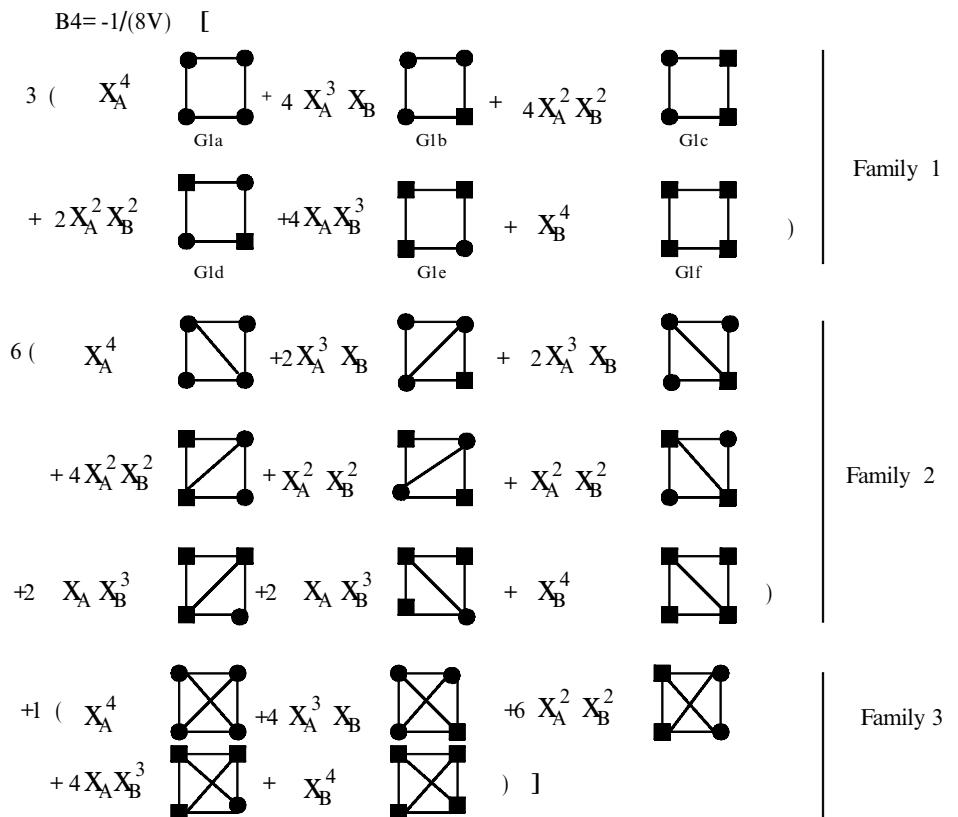
Figure 4. Graphs appearing in figure 3 arranged in a different way to allow comparison with the graphs of figure 2.

(c) Evaluate the integrals and the value of the parentheses of figure 2 for the choice made in step (b) of the chemical identity of molecules 1, 2, 3 and 4.

(d) Repeat steps (b) and (c) N_{chemical} times. Evaluate the average value of the parentheses of figure 2. Multiply that by $-1/(8V)$ thus obtaining the value of the fourth virial coefficient of the mixture for the chosen composition.

The reader may be surprised that the algorithm described provides the correct result. Let us illustrate, for a simple case, that this is so. For a binary mixture, the graphs appearing in the fourth virial coefficient are well known [18] and are shown in figure 3. Now black circles represent component A and black squares represent component B. To each filled symbol (circle or square) we associate the function 1. The graphs appearing in figure 3 can be rearranged into three

families denoted as family 1, family 2 and family 3 as shown in figure 4. Let us compare the graphs of family 1 of figure 4 with the graph G1 of figure 2. We shall now discuss what happens when starting from the graph G1 of figure 2 when we assign randomly the chemical identity of circles 1, 2, 3 and 4 with a probability x_A of being A and a probability x_B of being B. The probability of generating graph G1a starting from G1 is given by x_A^4. The probability of generating G1b is given by $4x_A^3 x_B$. This is so because molecule B may appear in any of the four vertices of G1. The probability of generating G1c is given by $4x_A^2 x_B^2$. This is so because the two black circles can occupy any of the four sides of the square. It is easy to show that G1d is generated with probability $2x_A^2 x_B^2$, G1e with probability $4x_A x_B^3$ and G1f with probability x_B^4. The sum of the probabilities of generating graphs G1a to G1e is 1 since they correspond to coefficients of the binomial expansion $(x_A + x_B)^4$ whose value is one. Therefore, starting from graph G1 of figure 2 and assigning randomly the identity of molecules 1, 2, 3 and 4 one obtains the graphs of family 1 of figure 4 with the correct probability. The same analysis can be performed with the graphs of family 2 of figure 4 and the graph G2 of figure 2 and for the graphs of family 3 of figure 4 and the graph G3 of figure 2. The conclusion is that by randomly generating the chemical identity of molecules 1, 2, 3 and 4 in figure 2 one obtains the graphs of figure 4 with the correct probability. In section 5 of this work we shall illustrate that the virial coefficients obtained from this procedure are coincident with those obtained previously by other authors for hard spheres mixtures.

Therefore the basic idea of this work is to use the graphs of figure 2 (i.e. the graphs of a pure fluid) for determining the virial coefficient of a mixture by choosing randomly (with probability given by the molar fraction) the chemical identity of the molecule assigned to each vertex of the diagram. Although we have illustrated the procedure for the fourth virial coefficient, it can be applied equally well to the second, third, fifth and the rest of the virial coefficients. In summary, use the graphs of a pure fluid, but choose randomly, with a probability given by the composition, the chemical identity of each vertex.

Now let us discuss briefly some of the advantages and disadvantages of the method proposed in this work.

Disadvantage 1. What happens if one wants to compute the virial coefficients of a binary mixture not just for a certain composition but for any composition? If one uses the classic methodology, for instance by using the graphs of figure 3 then one must evaluate only the graphs of each parenthesis and then one can obtain the virial coefficient for any other composition.

With the new method the virial coefficient is determined for a certain composition only. Therefore for other compositions it is necessary to repeat the calculations. However taking into account that the fourth virial coefficient of the binary mixture can be expressed as:

$$B_4 = \sum_{i=A,B} \sum_{j=A,B} \sum_{k=A,B} \sum_{l=A,B} x_i x_j x_k x_l B_4(ijkl)$$
$$= x_A^4 B_4(AAAA)$$
$$+ 4x_A^3 x_B B_4(AAAB) + 6x_A^2 x_B^2 B_4(AABB)$$
$$+ 4x_A x_B^3 B_4(ABBB) + x_B^4 B_4(BBBB). \qquad (8)$$

According to equation (8) if the fourth virial coefficient is known for five different compositions then the values of $B_4(AAAA)$, $B_4(AAAB)$, $B_4(AABB)$, $B_4(ABBB)$, $B_4(BBBB)$ can be determined by solving a linear system and after that the value of B_4 can be estimated for any other composition. For a three-component mixture however, the number of coefficients of the equation equivalent to equation (8) is 18, so one needs to know the virial coefficient for 18 different compositions in order to evaluate all of the coefficients. An alternative is to directly compute the individual coefficients of equation (8). This can be done easily. For instance the coefficient $B_4(AAAB)$ can be obtained by the following algorithm which will be denoted hereafter as Algorithm 1.

Algorithm 1

(a) Locate randomly three particles of type A and one of type B in the vertex labelled as 1,2,3,4 in figure 1.

(b) Evaluate the integrals (G1,G2,G3) and the value of the parentheses of figure 1 for the choice made in step (a).

(c) Repeat steps (b) and (c) a number of $N_{chemical}$ times. Evaluate the average value of the parenthesis of figure 1. Multiply that by $-1/(8V)$ thus obtaining the value of the fourth virial coefficient B_{AAAB}.

Disadvantage 2. From a numerical point of view the 'classic' procedure is more efficient computationally. In other words, longer calculations are required with the procedure described in this work to obtain the same numerical accuracy. This is so because in the classic procedure the composition is not a sampled variable, whereas in the procedure of this work it is a variable that must be sampled and that introduces certain statistical error (the chemical identity of molecules 1, 2, 3 and 4 must be selected randomly many times in order to mimic the imposed composition). For binary mixtures there is no doubt that the classic

$B4 = -1/(8V)$ [-2   +3   ]

$B4 = -1/(8V)$ [-2]
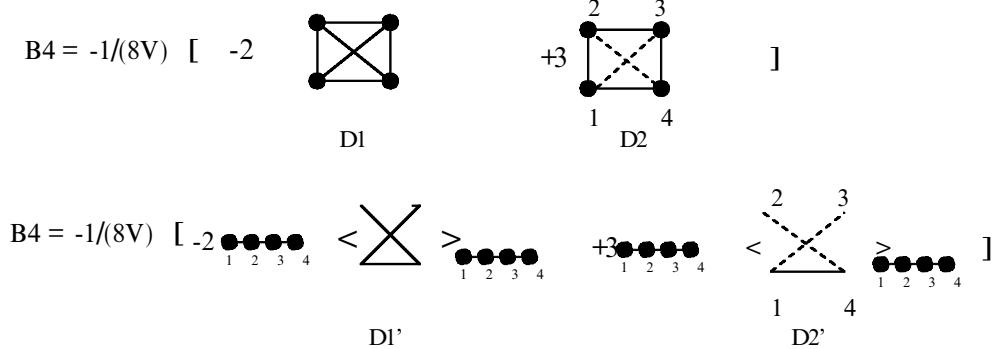
Figure 5. Ree and Hoover graphs required to evaluate B_4. Dashed lines indicate an e function ($e = \exp(-\beta u)$). The value of each diagram is obtained as the average of a certain function times the value of the chain diagram.

procedure is preferred for the lower virial coefficients (up to the fifth for which all the graphs are well known). However for higher virial coefficients, sixth and seventh, then the algorithm of this work may even be useful for binary mixtures.

Advantage 1. This new procedure is especially suitable for multicomponent mixtures. The graphs are still given by figure 2 and the only complication is that now the chemical identity of each vertex must be chosen among a higher number of components. The method proposed in this work is quite simple to implement, even when the number of components is large. The analogue of figure 3 for a multicomponent mixture has a higher number of graphs (i.e. it must include graphs with three and four different species). In the case of the fifth virial coefficient of a multicomponent mixture this is even more pronounced and the analogue of figure 3 will include a large number of graphs. This increases the probability of a mistake and the programming effort. In any case the new procedure can always be used as a cross-check since it is quite simple to implement.

Advantage 2. The new procedure can be used easily in the case of a system with an infinite number of components, for instance a polydisperse mixture of hard spheres, or a polymer which may adopt an infinite number of conformations each of which can be treated as a different component. In this case the 'classic' procedure cannot be implemented at all since one cannot evaluate the graphs for all possible combinations of the components since the number of components is now infinite.

Before finishing this section let us mention more precisely the way in which the fourth virial coefficients have been evaluated. Instead of using the graphs presented in figure 2 we shall use the graphs introduced by Ree and Hoover [4] for a pure fluid. They are presented in figure 5. Now a dashed line represents the Boltzmann factor of the potential between a pair of molecules. This

has the advantage that we must evaluate two instead of three different diagrams. We shall study in this work molecules with hard body intermolecular interactions, so that the Boltzmann factor takes the values zero or one and the Mayer function takes the value -1 or zero. For hard bodies Ree and Hoover have proposed a way of evaluating the diagrams of figure 5. In fact they showed that the integrals appearing in figure 5 can be written in terms of the average of a certain function over the chain configuration (the chain configuration is obtained when molecule 1 overlaps with molecule 2, molecule 2 overlaps with molecule 3 and molecule 3 overlaps with molecule 4). The value of the chain diagram is related to the product of the crossed second virial coefficient between molecules 1 and 2, 2 and 3, and 3 and 4. We shall describe now in full detail the algorithm to compute the fourth virial coefficient of hard sphere mixtures and of a hard polymer model.

3. Virial coefficients for hard sphere mixtures

The algorithm to compute the fourth virial coefficient of a hard sphere multicomponent mixture, which will be denoted hereinafter as Algorithm 2 (spheres), is as follows.

Algorithm 2 (spheres)

(a) Choose a composition for which you wish to evaluate the value of B_4.

(b) Randomly select the chemical identity (or state) of molecule 1 from the distribution $s(\alpha)$. For a multicomponent mixture this equates to choosing the identity of molecule 1 with probability x_A of being of type A. Repeat the procedure for molecules 2, 3 and 4.

(c) Fix molecule 1 on the origin of the coordinates of the system. Then randomly choose a position for molecule 2 so that molecule 1 overlaps with molecule 1. This is accomplished by selecting a point within a sphere of radius $L_{12} = (\sigma_1 + \sigma_2)/2$. This guarantees that molecule 2 overlaps with molecule 1.

(d) Move the origin of the laboratory system to molecule 2 and randomly select the position of molecule 3 so that molecule 3 overlaps with molecule 2 (using the same procedure described in point (c)). Once this is achieved move the origin to molecule 3, and select randomly the position of molecule 4 so that molecule 4 overlaps with molecule 3. The chain graph has been generated.

(e) Evaluate the functions $D1'$ and $D2'$ (see the functions in brackets in figure 5) for the generated chain configuration. Repeat steps (c), (d) and (e) N_{orien} times keeping the chemical identity of molecules 1, 2, 3 and 4. A good choice for N_{orien} is 10 000. Obtain the average value of $D1'$ and $D2'$ obtained over the N_{orien} configurations and multiply this average value by the value of the chain configuration, i.e. $\left(-2VB_2(1,2)\right)\left(-2B_2(2,3)\right)\left(-2B_2(3,4)\right)$ where:

$$B_2(1,2) = \frac{1}{2}\frac{\pi}{6}(\sigma_1 + \sigma_2)^3, \qquad (9)$$

with similar expressions for $B_2(2,3)$ and $B_2(3,4)$. Therefore the value of the diagram $D1$ of figure 5 for the chosen identity of molecules 1, 2, 3 and 4 is given by:

$$D1 = \langle D1' \rangle \left(-8VB_2(1,2)B_2(2,3)B_2(3,4)\right) \quad (10)$$

and the value of $D2$ is obtained from a similar expression.

(f) Repeat steps (b), (c), (d) and (e) N_{chemical} times. The optimal number of N_{chemical} can vary from one problem to another. N_{chemical} should guarantee that the chemical composition of the system is properly sampled. For binary mixtures a reasonable value for N_{chemical} is 5000. When the number of components increases then it is a good idea to increase N_{chemical} and to decrease N_{orien}.

(g) The value of B_4 for the chosen composition is just $-1/(8V)(-2\langle D1 \rangle + 3\langle D2 \rangle)$, where $\langle D1 \rangle$ is the average value of $D1$ over the N_{chemical} configurations.

For a polydisperse hard sphere mixture the algorithm described in steps (a)–(g) is still valid. The only difference is that step (a) is suppressed and in step (b) the chemical identity of molecules 1, 2, 3 and 4 is taken from the distribution function $s(\alpha)$.

4. Algorithm to compute virial coefficients of a polymer

The algorithm to compute the virial coefficient of a polymer molecule resembles that described in the pre-vious section for hard spheres. We shall describe a procedure to compute the fourth virial coefficient of a hard flexible molecule. Let us first mention the main differences between the polymer and the hard sphere case.

(i) The number of components, which is in principle finite for the hard sphere mixture (although not for the polydisperse hard sphere mixture), is infinite for a flexible molecule with continuous torsional angles. Within the RIS approximation, however, the number of possible configurations of the chain is huge but finite. In any case, the number of possible configurations (components) of the chain becomes quite a big number.

(ii) In addition to the configuration (conformation) of the molecule, now the molecule has orientational degrees of freedom.

(iii) Generation of chemical identities (conformations) from the probability distribution function $s(\alpha)$ becomes a difficult problem.

The last point can be solved by using a Monte Carlo simulation [39] prior to the determination of the virial coefficients. From this MC simulation a number of conformers will be selected from the distribution function $s(\alpha)$. The second point is solved by including orientational degrees of freedom within the determination of the virial coefficient. The first problem is probably the more serious one. It means that we cannot compute all possible contributions to the fourth virial coefficient. We will make a selection of chemical identities and hope that they constitute the main contribution to the studied virial coefficient. Let us recall at this point that this is a common problem in numerical simulation. For instance when computing the internal energy of a Lennard-Jones (LJ) fluid within a computer simulation, we are not exploring all the phase space available (configurations) of the system but just a small part and assume that the average of the internal energy obtained from our short trip through the phase space is representative of the average of the whole available phase space.

Therefore, the proposed algorithm to compute the fourth virial coefficient of a chain molecule, which will be denoted hereinafter as Algorithm 2 (flexible), is the following.

Algorithm 2 (flexible)

(a) Perform a Monte Carlo of an isolated chain. We shall use the pivoting algorithm to move the configuration of the molecule [39]. New configurations will be sampled with the Metropolis acceptance algorithm. Once the chain is equilibrated, about N_{photo} instantaneous configurations will be stored for later analysis. Since these N_{photo} configurations were generated

through a Monte Carlo simulation, one may think of this as representing a selection of N_{photo} chemical identities of the polymer molecule selected from the probability distribution $s(\alpha)$. A good value for N_{photo} would be 10 000. The value of N_{photo} should be high so that the probability of having two identical configurations of the polymer within the same graph becomes quite small. For long polymers, having two identical configurations of the chain within the same graph may bias the result.

(b) Select randomly one of the N_{photo} molecules obtained in step (a). This molecule will be the molecule labelled as 1 in the graph. Select randomly a second one from the N_{photo} molecules of step (a). This will be labelled as molecule 2. Repeat the procedure for molecules 3 and 4.

(c) Fix the centre of mass of molecule 1 in the origin of coordinates of the system and assign to molecule 1 a random orientation. To assign a random orientation to molecule 1 we generate randomly the three Euler angles which define the orientation of the molecule. We typically generated $\cos(\theta)$ uniformly between -1 and 1 and ϕ and γ uniformly between 0 and 2π.

(d) Choose randomly a position for the centre of mass and an orientation for molecule 2. The centre of mass of molecule 2 is chosen randomly within a sphere of radius L_{12} centred on the location of the centre of mass of molecule 1. The value of L_{12} is chosen so that overlapping between molecules 1 and 2 is impossible when


Molecule 1   Molecule 2
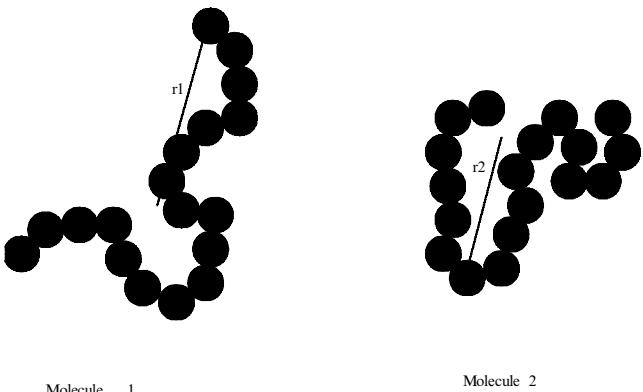
Figure 6. Choice of L_{12} for a pair of instantaneous configurations of the polymer. r_1 is the maximum of all the distances between the centre of mass of molecule 1 and the monomers of the polymer. r_2 is the maximum of all the distance between the centre of mass of molecule 2 and the monomers of the polymer. Notice that molecules 1 and 2 represent two different instantaneous configurations of the polymer. The choice made for L_{12} in this work is $L_{12} = r_1 + r_2 + \sigma$.

the distance between the centre of mass of molecules 1 and 2 is larger than L_{12} regardless of the orientations of molecules 1 and 2. The way in which L_{12} is selected is illustrated in figure 6. The random orientation of molecule 2 is obtained by selecting randomly the three Euler angles as described in step (c). Once the centre of mass location and orientation of molecule 2 has been determined we check for overlap between molecules 1 and 2. If there is no overlap a new location and orientation for molecule 2 is generated. The procedure is repeated until molecule 2 overlaps with molecule 1. We denote $N_{trial}(1,2)$ as the number of attempts required to generate a trial configuration in which molecule 2 overlaps with molecule 1.

(e) Move the origin of the laboratory system to molecule 2 and select randomly the position of the centre of mass and orientation of molecule 3 until molecule 3 overlaps with molecule 2 using the same procedure described in point (d). Let $N_{trial}(2,3)$ be the number of attempts required to generate a configuration where molecule 3 overlaps with molecule 2 and L_{23} the radius of the sphere where the trial positions of the centre of mass of molecule 3 were generated. Remember that overlaps between molecules 2 and 3 should not be possible for distances between the centre of mass larger than L_{23}.

(f) Move the origin of the laboratory system to molecule 3 and proceed as in step (e) for molecule 4. Once this is done, the chain graph (where molecule 1 overlaps with molecule 2, molecule 2 overlaps with molecule 3 and molecule 3 overlaps with molecule 4) has been generated.

(g) Evaluate the function $D1'$ and $D2'$ (which are plotted in figure 5) for the generated chain configuration. Repeat steps (c), (d), (e) and (f) N_{orien} times keeping the chemical identity of molecules 1, 2, 3 and 4. A good choice of N_{orien} for polymer molecules is 80 000. Obtain the average value of the functions $D1'$ and $D2'$ over the N_{orien} configurations and multiply this average value by the value of the chain configuration, i.e. $(-2VB_2(1,2))(-2B_2(2,3))(-2B_2(3,4))$. The crossed virial coefficients $B_2(1,2)$ are estimated as:

$$B_2(1,2) = \frac{1}{2}\frac{4\pi}{3}L_{12}^3 \frac{N_{orien}}{\displaystyle\sum_{k=1}^{k=N_{orien}} N_{trial\,k}(1,2)}. \qquad (11)$$

Table 1. Partial fourth virial coefficients for binary mixtures of hard spheres, with component A of diameter 1 and component B of diameter R. For each model the results of the first row are those of Enciso *et al.* [22] and were obtained by detailed evaluation of all the diagrams. The results of the second row were obtained with Algorithm 1 of this work. Results are given in units of σ_A^9. Results of this work were obtained with $N_{orien} = 20\,000$ and $N_{chemical} = 1000$.

R	$B_4(AAAA)$	$B_4(AAAB)$	$B_4(AABB)$	$B_4(ABBB)$
0.05	2.64(2)	0.0554(5)	0.431(5) 10^{-4}	0.179(3) 10^{-7}
0.05	2.63	0.0553	0.431 10^{-4}	0.173 10^{-7}
0.2	2.65(2)	0.159(1)	0.434(2) 10^{-2}	0.849(4) 10^{-4}
0.2	2.63	0.159	0.433 10^{-2}	0.849 10^{-4}
0.6	2.62(2)	0.888(6)	0.287(2)	0.892(4) 10^{-1}
0.6	2.63	0.898	0.289	0.893 10^{-1}
0.8503	2.65(2)	1.85(1)	1.28(1)	0.889
0.8503	2.63	1.84	1.28	0.889

Therefore the value of the diagrams $D1$ on the right hand side of figure 5 for the chosen identity of molecules 1, 2, 3 and 4 is given by:

$$D1 = \langle D1'\rangle\left(-8VB_2(1,2)B_2(2,3)B_2(3,4)\right) \quad (15)$$

and with a similar expression for $D2$.

(h) Repeat steps (b), (c), (d), (e), (f) and (g) a number of $N_{chemical}$ times. For polymer molecules we found $N_{chemical} = 1000$ to be a reasonable choice. For polymers steps (b) to (g) are computationally expensive so for calculations of the fourth virial coefficient the product of N_{orien} times $N_{chemical}$ should not exceed 10^8. The number of N_{photo} should be larger than $N_{chemical}$ by a factor of 4 or more.

(i) The value of B_4 is just $-1/(8V)$ $\left(-2\langle D1\rangle + 3\langle D2\rangle\right)$, where $\langle D1\rangle$ is the average value of $D1$ over the $N_{chemical}$ configurations.

Our choice of N_{photo}, N_{orien} and $N_{chemical}$ is orientative. It works reasonably well and it serves to illustrate that correct results are obtained from this algorithm. However it may not be optimum from a computational point of view so that further research is needed to improve numerical efficiency. Here our main interest is to show how to implement the algorithm and that it provides correct results. Let us finally mention an alternative possible way of implementing steps (a) and (b). In step (a) one could perform four independent Monte Carlo simulations of an isolated chain so that four different sets of N_{photo} instantaneous configurations are obtained, say set 1, set 2, set 3 and set 4. Then in step (b) one could choose randomly one of the sets, and one of the molecules within that set denoting the chosen molecule as molecule 1. Molecule 2 should be chosen randomly among the molecules of any of the three remaining sets. At this point two sets have been used and two not. Then choose molecule 3 randomly from

Table 2. Partial fourth virial coefficients $B_4(ABCD)$ and $B_4(AABC)$ for a four component hard sphere mixture with diameters 1, 0.8, 0.6 and 0.4 for components A, B, C and D respectively. Results are given in units of $\left(\pi\sigma_A^3/6\right)^3$. When implementing Algorithm 1 we used $N_{orien} = 40\,000$ and $N_{chemical} = 10\,000$. The results of the first row correspond to the numerical determination of this work whereas the results of the second row correspond to the estimate obtained from the BMCSL equation of state.

$B_4(ABCD)$	$B_4(AABC)$
0.530	3.754
0.519	3.677

any of the two non-used sets, and finally choose molecule 4 randomly from the non-used set. In this way it is very unlikely that the same configuration of the polymer appears twice or more in any of the graphs.

5. Results

In this section the results of this work will be presented. First we focus on hard sphere binary mixtures. The diameter of the first component, i.e. component A, will be assigned to one and the diameter of the second component, i.e. component B, will be R (with R less than one). In table 1 the individual virial coefficients of a binary mixture (see equation (8)) as reported by Enciso *et al.* [22] as obtained in this work from the algorithm denoted as Algorithm 1 are presented for several size ratios. As can be seen the results of this work agree quite well with those of Enciso *et al.* [22] and are in all cases within the statistical uncertainty reported by Enciso *et al.* One can conclude from the results of table 1 that Algorithm 1 presented in this work is a simple way of evaluating the individual virial coefficients $B_4(AAAA)$, $B_4(AAAB)$, $B_4(AABB)$, $B_4(ABBB)$,

Table 3. Fourth virial coefficient for an equimolar binary mixture ($x_A = 0.5$) of hard spheres of diameter 1 for component A and diameter R for component B. Results of the column labelled as Enciso *et al.* were obtained by taking the results of [22] for the individual virial coefficients and using equation (8) for determining the virial coefficient of the mixture. Results labelled as Algorithm 1 correspond to the results of this work by using equation (8) and the algorithm denoted as Algorithm 1 to compute the individual virial coefficients. Results labelled as Algorithm 2 were obtained directly for the mixture by using the algorithm of this work denoted as Algorithm 2. We use $N_{orien} = 40\,000$ and $N_{chemical} = 2000$ for Algorithm 2. Results are given in units of σ_A^9.

R	Enciso *et al.*	Algorithm 1	Algorithm 2
0.05	0.179	0.178	0.177
0.2	0.207	0.206	0.204
0.6	0.517	0.521	0.517
0.8503	1.369	1.365	1.360

Table 4. Second, third and fourth virial coefficient of a multicomponent mixture of hard spheres. The mixture presents ten different components with diameters 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2 and 0.1 being 1 the diameter of component A. The molar fraction of each component of the mixture is 0.10. The virial coefficients are given in reduced units, so that $B_2^* = B_2/(\pi\sigma_A^3/6)$, $B_3^* = B_3/(\pi\sigma_A^3/6)^2$, $B_4^* = B_4/(\pi\sigma_A^3/6)^3$. The results of the first row corresponds to the numerical results of this work using Algorithm 2 with $N_{orien} = 30\,000$ and $N_{chemical} = 30\,000$. The results of the second row correspond to the estimate of the virial coefficients from the Carnahan–Starling equation of state of hard sphere mixtures.

B_2^*	B_3^*	B_4^*
0.937	0.644	0.344
0.938	0.647	0.340

$B_4(BBBB)$ of a binary mixture. It should be noted that algorithm 1 of this work is not limited to binary mixtures. In table 2 we show some individual coefficients for a mixture of four components denoted as A, B, C and D and with diameters 1, 0.8, 0.6 and 0.4, respectively. For this mixture of four components, individual fourth virial coefficients have never been reported so that the results of table 2 are completely new. We also show in table 2 the virial coefficients obtained from the Boublik–Mansoori–Carnahan–Starling–Leland (BMCSL) equation of state (EOS) for hard sphere mixtures [40, 41]. These virial coefficients are obtained from the corresponding derivatives of the equation of state [42] and using the mathematical package MAPLE [43] for performing the algebraic manipulations. The agreement between theory and numerical results is quite good. The results of table 2 illustrate how the algorithm, denoted as Algorithm 1, can be applied to compute the individual virial coefficients of any multicomponent mixture.

Although for multicomponent mixtures one may be interested in individual virial coefficients like for instance $B_4(ABCD)$, one is usually interested in the value of the fourth virial coefficient for a certain composition. In table 3 the fourth virial coefficient, as defined by equation (1) for an equimolar binary mixture of hard spheres, is presented. In the second and third column the results obtained by using equation (8) and the individual virial coefficients presented in table 1 are shown. Again the agreement is good as one would expect from the high agreement already shown in table 1. More interesting are the results of the fourth column. These results were obtained by using Algorithm 2 (spheres) of this work. Let us recall that in Algorithm 2 (spheres) we do not compute the indi-

vidual virial coefficients $B_{AABB}...$ but we simply attempt to compute the global value of B_4 for the mixture. As can be seen the results of Algorithm 2 (spheres) are in close agreement with those obtained from equation (8) with the individual virial coefficients obtained either from Algorithm 1 or from the exhaustive enumeration of diagrams of Enciso *et al.* The results of table 3 show that Algorithm 2 (spheres) can be used with confidence to evaluate the global value of B_4 of a multicomponent mixture of a certain composition. Let us show this for a more interesting case. Let us consider a 10-component mixture of hard spheres with molar fraction 0.10 for all the components and with diameters 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2 and 0.1. To the best of our knowledge the fourth virial coefficient has never been reported for a 10-component mixture. One may suspect that the process of colouring the diagrams presented in figure 1 for a 10-component mixture is tedious and the number of diagrams resulting from this process is quite big. In contrast Algorithm 2 (spheres) of this work is quite simple to implement. The result of B_2, B_3 and B_4 of this mixture are presented in table 4. In this table the virial coefficients predicted by the BMCSL equation of state for this mixture are also presented. As can be seen the agreement between the numerical results of this work and the prediction of the theoretical equation of state is quite good. That constitutes a cross-check of the calculations of this work and also confirm the accuracy of the BMCSL EOS for this system. The results of table 4 illustrate how virial coefficients are readily obtained for a multicomponent mixture.

Let us now change to mixtures with either infinite or finite, but huge, number of components (as for flexible molecules). In figure 7 the model for the polymer considered in this work is shown. The model is related (but is is not identical) to the so-called 'pearl-necklace' chain.
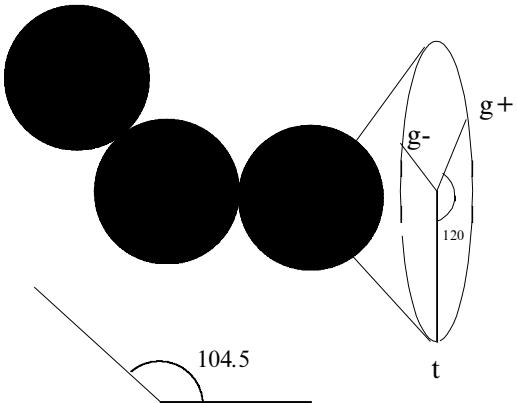
Figure 7. Model for the polymer studied in this work. The polymer is formed by tangent hard spheres. The angle between three monomer units is fixed to 104.5°. Each torsional angle may adopt three torsional states denoted as trans, gauche+ and gauche−. We shall assume that the energy of these three torsional states is zero. Monomers within a chain separated by four or more bonds interact through a hard sphere intramolecular potential (i.e. we have a HS interaction between monomers 1 and 5). Interaction between a pair of molecules is given by a hard sphere site–site potential.

Table 5. Second, third and fourth virial coefficient as defined by equation (1) for the model of the polymer described in figure 7. The number of monomer units of the monomer is denoted as m. We used $N_{orien} = 80\,000$ and $N_{chemical} = 1000$. Virial coefficients are given in terms of the molecular volume V_m which is given by $V_m = \pi m \sigma^3/6$, where σ is the diameter of the hard sphere associated with each monomer unit.

m	B_2/V_m	B_3/V_m^2	B_4/V_m^3
7	10.25	55.3	160
10	12.51	79.4	246
16	16.64	133.6	444
32	25.80	301.4	1049

The polymer is formed by m tangent hard spheres, with the bond angle fixed to 104.5°, with three torsional states denoted as trans (t), gauche+ (g^+) and gauche− (g^-) having the same energy (i.e. zero). Intramolecular interactions between pairs of monomers separated by four or more bonds and intermolecular interactions are of hard sphere (HS) type. In table 5 the second, third and fourth virial coefficient for this simple model of polymer are presented for lengths up to $m = 32$. Obviously the computing time for determining virial coefficients increases strongly with m. For the largest value of m calculations were performed in a parallel machine (Seymour 2000) using ten processors simultaneously. The calculation of B_4 would have required about 12 days of CPU in a
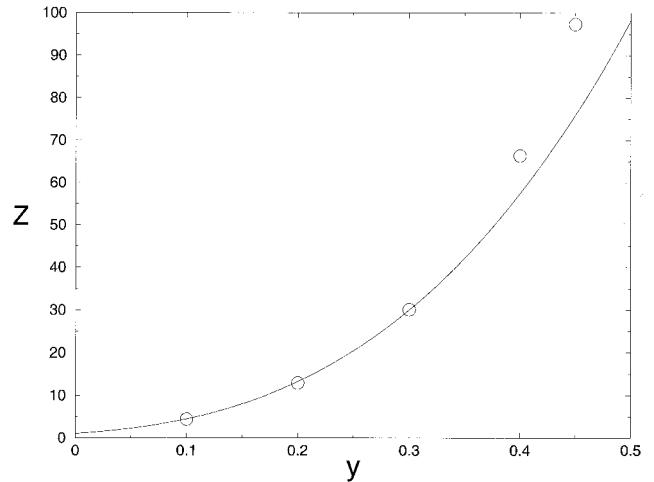

Figure 8. Compressibility factor at low and moderate densities of the fully flexible pearl necklace model of $m = 16$ tangent hard spheres as obtained from the Monte Carlo simulations of [44] (symbols) and from the virial expansion of equation (1) truncated at B_4 for the model of the polymer described in figure 7. The volume fraction y is defined as $y = \rho V_m$, where the molecular volume V_m is given by $V_m = \pi m \sigma^3/6$.
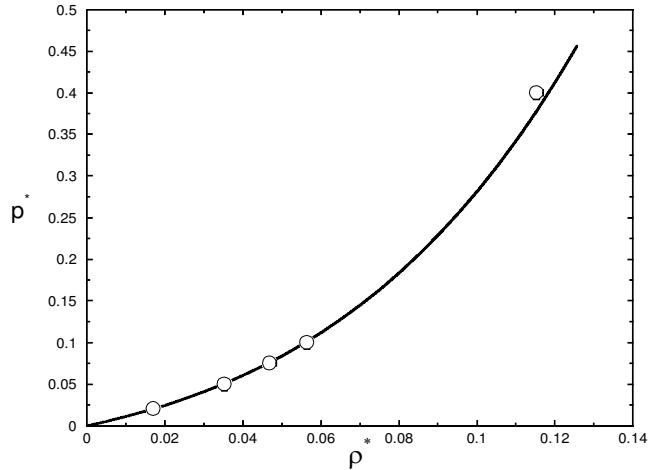

Figure 9. Equation of state at low and moderate densities of a hard n-alkane model of n-hexane. Results were obtained at $T = 300$ K. Pressure and density are given by reduced units (i.e. $p^* = p/(kT/\sigma^3)$ and $\rho^* = \rho\sigma^3$). Solid line: virial expansion truncated at the fourth virial coefficient. Symbols: simulation results [49].

single processor conventional workstation. In figure 8 the compressibility factor, as obtained from computer simulations of the pearl-necklace model [44], are compared with the results of the virial expansion truncated in B_4, with the virial coefficients obtained in this work for the model with $m = 16$. Although the model for the polymer used in this work is not identical to the fully flexible pearl-necklace model one expects only small differences in the virial coefficients between these two

models (in fact within Wertheim's theory they should have the same equation of state [45, 46]). As can be seen in figure 8 the EOS obtained from simulations agree quite well with the results of the virial expansion of this work providing some indirect evidence for the correctness of the calculations.

As another example we have evaluated the second, third and fourth virial coefficients of a hard n-alkane model of n-hexane (see the M3 model of a recent paper by Malanoski and Monson [47] for more details). In the model the bond length is fixed to $L^* = L/\sigma = 0.40$. The bond angle is fixed to the tetrahedral value $109.5°$. The torsional potential between the pair of atoms separated by three bonds is that proposed by Ryckaert and Bellemans [48]. The pair potential between the pair of atoms separated by four or more bonds and between atoms in different molecules is a hard sphere site–site potential with diameter σ. Due to the torsional potential the properties of the model depend on temperature. Results were obtained for $T = 300$ K. Since the torsional potential is continuous the number of possible configurations of the chain is infinite (i.e. each of the three torsional angles of the molecule can adopt any value between zero and 2π). The obtained value of the second, third and fourth virial coefficient are $B_2^* = B_2/\sigma^3 = 10.05$, $B_3^* = B_3/\sigma^6 = 58.4$ and $B_4^* = B_4/\sigma^9 = 224$, respectively. In figure 9 the equation of state obtained from the truncated virial expansion (truncated at B_4) is compared to the simulation data obtained at low and moderate densities [49]. As can be seen the agreement is quite good providing, again, an indirect indication of the correctness of the calculations for the virial coefficients.

6. Conclusions

In this paper we have proposed a methodology to compute virial coefficients of multicomponent mixtures. The methodology is useful regardless of the number of components of the mixture so it can be applied to mixtures containing from two to an infinite number of components. We have presented results for the virial coefficients up to the fourth for mixtures of hard spheres with four or even ten different components. The agreement with the predictions of the BMCSL equation of state is good for the considered mixtures. We also present results for the virial coefficients up to the fourth of a polymer model with 32 monomer units and for a hard model of n-hexane. To the best of our knowledge this is the first time that the fourth virial coefficient has been reported for a flexible molecule, and the first time that the third virial coefficient has been reported for a non-lattice flexible molecule. The proposed algorithm is quite simple to implement especially if a code for evaluating the virial coefficients of a pure non-spherical fluid is

available. This is so because we use the same graphs for the multicomponent system as those for the pure fluid so that the programming effort to extend the pure fluid code to the multicomponent mixture is small. We hope the methodology described here will encourage further studies on virial coefficients of multicomponent hard sphere mixtures and especially for polymer molecules for which very little is known.

References

[1] MAYER, J. E., and MAYER, M. G., 1940, *Statistical Mechanics* (New York: Wiley).

[2] McQUARRIE, D. A., 1976, *Statistical Mechanics* (New York: Harper and Row).

[3] HANSEN, J. P., and McDONALD, I. R., 1986, *Theory of Simple Liquids*, 2nd Edn (New York: Academic Press).

[4] REE, F. H., and HOOVER, W. G., 1964, *J. chem. Phys.*, **40**, 939.

[5] ISIHARA, A., 1951, *J. chem. Phys.*, **19**, 397.

[6] MONSON, P. A., and RIGBY, M., 1978, *Molec. Phys.*, **35**,

[7] FRENKEL, D., 1988, *J. phys. Chem.*, **92**, 5314.

[8] ALLEN, M. P., EVANS, G. T., FRENKEL, D., and MULDER, B., 1993, *Adv. chem. Phys.*, **86**, 1.

[9] RIGBY, M., 1989, *Molec. Phys.*, **66**, 1261.

[10] RIGBY, M., 1970, *J. chem. Phys.*, **53**, 1021.

[11] SINDELKA, M., and BOUBLIK, T., 1999, *Molec. Phys.*, **96**, 243.

[12] BOUBLIK, T., and NEZBEDA, I., 1986, *Collec. Czech. chem. Commun.*, **51**, 2301.

[13] VEGA, C., 1997, *Molec. Phys.*, **92**, 651.

[14] VEGA, C., GARZON, B., and LAGO, S., 1994, *Molec. Phys.*, **82**, 1233.

[15] BARKER, J. A., and MONAGHAN, J. D., 1962, *J. chem. Phys.*, **36**, 2564.

[16] SUN, J., and TEJA, A. S., 1996, *J. phys. Chem.*, **100**, 17365.

[17] RIGBY, M., and SMITH, E. B., 1963, *Trans. Faraday Soc.*, **59**, 2469.

[18] SAIJA, F., FIUMARA, G., and GIAQUINTA, P. V., 1996, *Molec. Phys.*, **87**, 991.

[19] SAIJA, F., FIUMARA, G., and GIAQUINTA, P. V., 1996, *Molec. Phys.*, **89**, 1181.

[20] SAIJA, F., FIUMARA, G., and GIAQUINTA, P. V., 1997, *Molec. Phys.*, **90**, 679.

[21] WHEATLEY, R. J., SAIJA, F., and GIAQUINTA, P. V., 1998, *Molec. Phys.*, **94**, 877.

[22] ENCISO, E., ALMARZA, N. G., GONZALEZ, M. A., and BERMEJO, F. J., 1998, *Phys. Rev.* E, **57**, 4486.

[23] WHEATLEY, R. J., 1998, *Molec. Phys.*, **93**, 675.

[24] WHEATLEY, R. J., 1999, *Molec. Phys.*, **96**, 1805.

[25] BIBEN, T., and HANSEN, J. P., 1991, *Phys. Rev. Lett.*, **66**, 2215.

[26] DIJKSTRA, M., VAN ROIJ, R., and EVANS, R., 1999, *Phys. Rev. Lett.*, **82**, 117.

[27] DIJKSTRA, M., VAN ROIJ, R., and EVANS, R., 1999, *Phys. Rev. E*, **59**, 5744.

[28] ALMARZA, N. G., and ENCISO, E., 1999, *Phys. Rev. E*, **59**, 4426.

[29] BORSTNIK, B., 1992, *Vestn. Slov. Kem. Drus.*, **39**, 145.

[30] BORSTNIK, B., 1994, *Croat. Chem. Acta*, **67**, 143.

[31] FLORY, J. P., 1969, *Statistical Mechanics of Chain Molecules* (New York: Wiley).

[32] VEGA, C., GARZON, B., and LAGO, S., 1994, *J. chem. Phys.*, **100**, 2182.

[33] YETHIRAJ, A., HONNELL, K. G., and HALL, C. K., 1992, *Macromolecules*, **25**, 3979.

[34] WICHERT, J. M., and HALL, C. K., 1994, *Macromolecules*, **27**, 2744.

[35] RUBIO, A. M., and FREIRE, J. J., 1996, *Macromolecules*, **29**, 6946.

[36] BRUNS, W., 1996, *Macromolecules*, **29**, 2641.

[37] MACDOWELL, L. G., and VEGA, C., 1998, *J. chem. Phys.*, **109**, 5670.

[38] BRUNS, W., 1997, *Macromolecules*, **30**, 4429.

[39] ALLEN, M. P., and TILDESLEY, D. J., 1987, *Computer Simulation of Liquids* (Oxford: Clarendon Press).

[40] BOUBLIK, T., 1970, *J. chem. Phys.*, **53**, 471.

[41] MANSOORI, G. A., CARNAHAN, N. F., STARLING, K. E., and LELAND, T. W., 1971, *J. chem. Phys.*, **54**, 1523.

[42] ENCISO, E., and ALMARZA, N. G., 1998, *Molec. Phys.*, **95**, 635.

[43] MAPLE is a registered trade mark of Waterloo Maple Software, Waterloo, ON, Canada.

[44] ZHOU, Y., SMITH, S. W., and HALL, C. K., 1995, *Molec. Phys.*, **86**, 1157.

[45] WERTHEIM, M. S., 1987, *J. chem. Phys.*, **87**, 7323.

[46] CHAPMAN, W. G., JACKSON, G., and GUBBINS, K. E., 1988, *Molec. Phys.*, **65**, 1057.

[47] MALANOSKI, A. P., and MONSON, P. A., 1999, *J. chem. Phys.*, **110**, 664.

[48] RYCKAERT, J. P., and BELLEMANS, A., 1978, *Faraday Discuss. chem. Soc.*, **66**, 95.

[49] MACDOWELL, L. G., 2000, private communication.