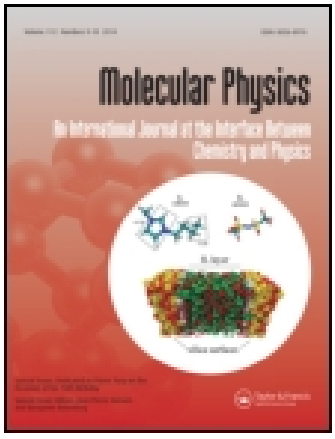

Molecular Physics: An International Journal at the Interface Between Chemistry and Physics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tmph20

Computation of high-order virial coefficients in high-dimensional hard-sphere fluids by Mayer sampling

Cheng Zhang[a] & B. Montgomery Pettitt[a]
[a] Sealy Center for Structural Biology and Molecular Biophysics, Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA
Published online: 23 Apr 2014.

To cite this article: Cheng Zhang & B. Montgomery Pettitt (2014) Computation of high-order virial coefficients in high-dimensional hard-sphere fluids by Mayer sampling, Molecular Physics: An International Journal at the Interface Between Chemistry and Physics, 112:9-10, 1427-1447, DOI: 10.1080/00268976.2014.904945

To link to this article: http://dx.doi.org/10.1080/00268976.2014.904945

Molecular Physics, 2014
Vol. 112, No. 9–10, 1427–1447, http://dx.doi.org/10.1080/00268976.2014.904945

INVITED ARTICLE

Computation of high-order virial coefficients in high-dimensional hard-sphere fluids by Mayer sampling

Cheng Zhang and B. Montgomery Pettitt*

Sealy Center for Structural Biology and Molecular Biophysics, Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA

(*Received 7 February 2014; accepted 5 March 2014*)

The Mayer sampling method was used to compute the virial coefficients of high-dimensional hard-sphere fluids. The first 64 virial coefficients for dimensions $12 < D \leq 100$ were obtained to high precision, and several lower dimensional virial coefficients were computed. The radii of convergence of the virial series in 13, 15, 17 and 19 dimensions agreed well with the analytical results from the Percus–Yevick closure.

Keywords: liquid state; equation of state

1. Introduction

For a simple one-component liquid, the virial coefficients, B_n, are the coefficients in the density, ρ, expansion of the pressure, P [1–3]:

$$\beta P = \sum_{n=1}^{+\infty} B_n \rho^n, \tag{1}$$

where $\beta = 1/(k_B T)$ is the inverse temperature, and $B_1 = 1$ from the idea gas limit $\rho \to 0$. The virial coefficient B_n is equal to a sum of integrals, each of which can be represented diagrammatically as a doubly connected or biconnected cluster of n field points [1]. Here we consider the virial coefficients of the hard-sphere fluid in high dimensions.

Hard-sphere fluids are simple model fluids in which spherical particles of unit diameter interact via only a hard-core repulsive potential: $u(r) = \infty$ for $r < 1$, or 0 otherwise for two particles separated by r. They are one of the oldest models for fluids and have served as a computational test ground for many theories [4] and methods [5,6]. Hard spheres allow a variety of mathematical and computational techniques not always applicable to soft potentials. Yet, one can argue that the insights achievable in this simple system are a useful guide for the studies of more realistic fluids [1].

Computing the virial coefficients of hard-sphere fluids is an old problem [7,8,9–11]. While B_2 and B_3 are known [3], and B_4 is known up to dimension $D = 12$ [7,11–13], higher order and/or higher dimensional coefficients are usually obtained from the Monte Carlo (MC) integration method pioneered by Metropolis *et al*. [5], and later refined by others [14–18]. The MC Integration method has been successfully applied to compute low-order virial coefficients in various dimensions [5,14,16–21]. Some notable achievements include those by Clisby and McCoy who obtained the virial coefficients up to $n = 6$ for $D \leq 50$ [20], and those up to $n = 10$ for $D \leq 8$ [16]. More recently, Wheatley computed the 11th and 12th virial coefficients for the three-dimensional case [18].

Although effective in two and three dimensions, the MC integration method is less powerful in higher dimensions. The main difficulty is that in high dimensions, the method rarely generates the biconnected clusters [22] needed by the virial expansion, as shown in Figure 1. This problem can be solved by replacing the MC integration method with a different importance sampling method.

The method used here is a special case of the Mayer sampling method [23,24], which was previously used to compute bridge diagrams and virial coefficients [24–26]. The idea is as follows. The sum of the cluster integrals of a certain order n in the Mayer expansion is compared to an analytically known ring integral [24], and the ratio of the two is measured from the corresponding averages [24] using importance sampling over the biconnected clusters. The method is particularly useful in high dimensions, for the ring integral dominates the cluster integrals there. Thus, the above ratio is close to unity, and can be computed to high precision. If, however, the average for the ring integral is difficult to obtain, we can replace the ring integral by the biconnected partition function, which can be obtained from a modified grand canonical ensemble simulation.

We develop the method and its extensions in Section 2. The results, including the first 64 virial coefficients for dimensions $12 < D \leq 100$, are presented in Section 3. We conclude the article in Section 4.

* Corresponding author. Email: mpettitt@utmb.edu

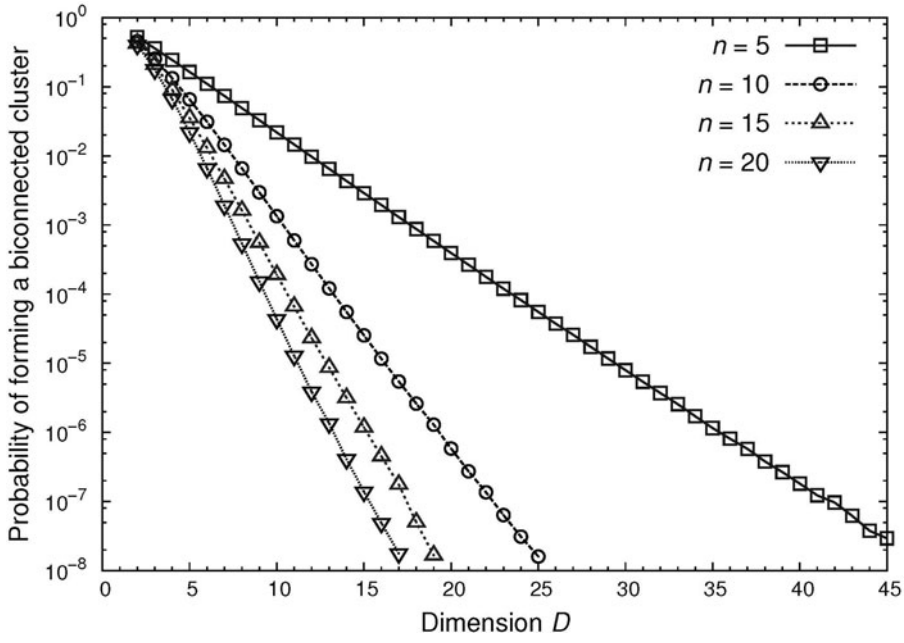

Figure 1. Probability of forming a biconnected configuration of n vertices from a linear chain generated from the MC integration method.

2. Methods

In this section, we describe the importance sampling method sketched in the Introduction along with some extensions. We first review the Mayer and Ree–Hoover (RH) expansions in Sections 2.1 and 2.2, as well as the biconnected partition function in Section 2.3. Next, we describe several methods of computing the RH star content in Section 2.4. We then discuss in Section 2.5 the conventional MC integration method and its difficulty in achieving the required biconnectivity. We define in Section 2.6 a ring content analogous to the star content, and use it to develop a basic Metropolis algorithm in Section 2.7. Sections 2.8 and 2.9 are devoted to a grand ensemble technique, allowing one to obtain the biconnected partition function and to compute a range of virial coefficients in a single simulation. Finally, in Section 2.10, we give two formulae to estimate the radius of convergence of the series (1) from the virial coefficients.

2.1. Mayer expansion

The Mayer expansion expresses the nth virial coefficient B_n as a sum of a few n-point configuration integrals:

$$B_n = \frac{1-n}{n!} \sum_k I^{(\mathrm{M})}_{n,k}, \tag{2}$$

where each integral $I^{(\mathrm{M})}_{n,k}$ is of the form of $\int g_{n,k}(\mathbf{r}^n)\, d\mathbf{r}^n$, and $\mathbf{r}^n = \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ are the coordinates of n D-dimensional points. Here $d\mathbf{r}^n \equiv d\mathbf{r}_2 \cdots d\mathbf{r}_n$, and $d\mathbf{r}_1$ is omitted to make the integral convergent. Instead of specifying $g_{n,k}(\mathbf{r}^n)$ algebraically, we represent each integral as a *Mayer diagram*, which is a simple graph of n vertices, e.g., $I^{(\mathrm{M})}_{3,1} = \triangle$, $I^{(\mathrm{M})}_{4,1} = \square$, …, $I^{(\mathrm{M})}_{4,10} = \boxtimes$. Below, we will adopt the standard terminology in graph theory, e.g., two vertices (edges) are *adjacent* if they share an edge (a vertex), and the *degree* of a vertex i is the number of vertices adjacent to i, etc.

For our purpose, a diagram is evaluated as follows. An edge between two vertices i and j carries an algebraic value of $f_{ij} \equiv \exp[-\beta\, u(r_{ij})] - 1$, where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, and $u(r_{ij})$ is the pair potential. The integrand $g_{n,k}$ is then the product of the f_{ij} values on all edges $\{i, j\}$ in the diagram [for it is derived from the expansion of the Boltzmann weight $\prod(1 + f_{ij})$], e.g., $\triangle = \int f_{12} f_{23} f_{31} d\mathbf{r}^3$, $\square = \int f_{12} f_{23} f_{34} f_{41} d\mathbf{r}^4$, and $\boxtimes = \int f_{12} f_{13} f_{14} f_{23} f_{24} f_{34} d\mathbf{r}^4$. Diagrams are labelled: we have implicitly placed the vertices on a circle in a counterclockwise fashion (and numbered, say, the vertex at the bottom-left corner as 1). The expansion also includes the diagrams $\bowtie = \int f_{13} f_{32} f_{24} f_{41} d\mathbf{r}^4$ and $\boxtimes = \int f_{12} f_{24} f_{43} f_{31} d\mathbf{r}^4$, although the two share the same value as $\square$. Under this convention every diagram has a unit coefficient aside from the overall factor $(1-n)/n\,!$ in Equation (2).

The virial expansion contains all and only biconnected diagrams [1]. To review, a diagram is *connected* if there is a path connecting any two vertices. Similarly, a connected diagram is doubly connected or *biconnected* (also called *irreducible*) if it has no *articulation point*, whose removal along with the removal of its incident edges breaks the diagram into parts. Similarly, a *separation pair* is a pair of vertices whose removal disconnects a connected diagram.

The Mayer expansion is somewhat inconvenient for MC calculations, in which we randomly generate configurations $\mathbf{r}^n = \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ and measure their contributions to the integrals $\int g_{n,k}(\mathbf{r}^n)\, d\mathbf{r}^n$. The Mayer diagrams are non-exclusive in that a configuration $\mathbf{r}^n$ may contribute non-zero $g_{n,k}(\mathbf{r}^n)$ values to multiple Mayer diagrams [16], e.g., the fully collapsed configuration $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}_3 = \mathbf{r}_4$ contributes to all diagrams □, ⧖, ⋈, ◰, ◱, ⊠, ... of four vertices. The sum over diagrams is avoided in the RH reformulation below.

2.2. *Ree–Hoover expansion*

The reformulated expansion uses two types of edges [14]. We call the edges in the Mayer expansion *f*-bonds, and define a new *e*-bond (also called $\tilde{f}$-bond) as $e_{ij} \equiv \exp[-\beta u(r_{ij})] = f_{ij} + 1$ for a pair $\{i, j\}$. Then, we expand a non-adjacent pair $\{i, j\}$ as the difference between the *e*-bond and the *f*-bond ($1 = e_{ij} - f_{ij}$), resulting in two diagrams: one carries the *e*-bond and the other the *f*-bond. Doing this to all non-adjacent pairs in all Mayer diagrams yields a new expansion in RH diagrams, in which every pair of vertices are joined by either an *f*-bond or an *e*-bond. Formally, we write

$$B_n = \frac{1-n}{n!} \sum_k S_{n,k} I^{(\mathrm{RH})}_{n,k}, \tag{3}$$

where the coefficient $S_{n,k}$ before each diagram or integral $I^{(\mathrm{RH})}_{n,k}$ is called the RH star content, which is not necessarily unity (here 'star' means a biconnected diagram). We will use the same symbols, e.g., △, □, ⊠,..., to draw RH diagrams; however, the missing edges are understood as *e*-bonds. The integrals are modified accordingly, e.g., $\square^{(\mathrm{RH})} = \int f_{12} f_{23} f_{34} f_{41} e_{13} e_{24} d\mathbf{r}^4$. The concepts of degree, connectivity, biconnectivity, etc., in an RH diagram are redefined in terms of the *f*-bonds.

We can compute the star content $S_{n,k}$ in Equation (3) as follows. First, we say that a Mayer or RH diagram H is a *subgraph* of an RH diagram G, if G contains all *f*-bonds in H. Now consider the transition from the Mayer expansion to the RH one. Because an RH diagram inherits all *f*-bonds from the original Mayer diagram, the latter is always a subgraph of the former. It follows that the virial expansion contains only biconnected RH diagrams [16]. Further, since each Mayer diagram H contributes at most once to a particular RH diagram G with a numeric value +1 or –1 (corresponding to an even or odd number of additional *f*-bonds, respectively), the star content of an RH diagram G is equal to the number of biconnected Mayer subgraphs H of G with an even number (including 0) of missing *f*-bonds, less the number of biconnected Mayer subgraphs with an odd number of missing *f*-bonds [15]. For example, the star content of ⊠ is $-2 = 1 + 3 - 6$ for there are three biconnected subgraphs □, ⧖, and ⋈ with two missing *f*-bonds, and six equivalent subgraphs, such as ◰ and ◱, with one missing *f*-bond. We refer to this result as the *RH theorem* below. More generally, if we wish to obtain an equivalent RH expansion of a restricted Mayer sum (e.g., of only the ring diagrams), then, in computing the coefficient before an RH diagram G, we must impose the same restriction on the subgraphs of G in using the RH theorem.

The RH expansion is suitable for a hard-sphere fluid because any configuration now contributes to only one RH diagram. Since the hard-sphere potential implies that f_{ij} is either -1 or 0, and e_{ij} is either 0 or 1, the necessary domain of integration of an RH diagram is restricted on every pair $\{i, j\}$. The integrand is non-zero if and only if $r_{ij} < 1$ for every *f*-bonded pair $\{i, j\}$, and $r_{ij} \geq 1$ for every *e*-bonded one. In this case, the ~~integral~~ is equal to $\prod_{r_{ij}<1} f_{ij} \prod_{r_{kl}\geq 1} e_{kl} = (-1)^{N_f}$, where N_f is the number of *f*-bonds in the diagram. Since $f_{ij} e_{ij} = 0$, any two RH diagrams have mutually exclusive domains of integration, and any configuration $\mathbf{r}^n$ is compatible with only one RH diagram [16]. We call this diagram the intrinsic RH diagram $G(\mathbf{r}^n)$ of the configuration $\mathbf{r}^n$. If $G(\mathbf{r}^n)$ is biconnected, we call $\mathbf{r}^n$ a *biconnected configuration*. Now, the sum over diagrams, or the corresponding integrals, in the RH expansion Equation (3) can be aggregated to a single integral over the biconnected configurations:

$$B_n = \frac{1-n}{n!} \int \mathrm{bc}(\mathbf{r}^n)\, S^*(\mathbf{r}^n)\, d\mathbf{r}^n, \tag{4}$$

where the function $\mathrm{bc}(\mathbf{r}^n)$ is 1.0 if $\mathbf{r}^n$ is biconnected or 0 otherwise. We call the quantity $S^*(\mathbf{r}^n) \equiv (-)^{N_f(\mathbf{r}^n)} S(\mathbf{r}^n)$ the *signed star content*, where $S(\mathbf{r}^n)$ and $N_f(\mathbf{r}^n)$ are the star content and the number of *f*-bonds of $G(\mathbf{r}^n)$, respectively. Comparing to Equation (2), we see that $S^*(\mathbf{r}^n)$ is equal to the sum of the integrands of the biconnected Mayer diagrams: $S^*(\mathbf{r}^n) = \sum_k g_{n,k}(\mathbf{r}^n)$, i.e., $S^*(\mathbf{r}^n)$ gives the overall contribution from $\mathbf{r}^n$ to the virial expansion.

2.3. *Biconnected partition function*

For later convenience, we define the biconnected partition function, or simply partition function, as

$$Z_n \equiv \int \mathrm{bc}(\mathbf{r}^n)\, d\mathbf{r}^n, \tag{5}$$

which can be rewritten as a sum over the biconnected RH diagrams:

$$Z_n = \sum_k \left| I^{(\mathrm{RH})}_{n,k} \right| = \sum_k (-1)^{N_f} I^{(\mathrm{RH})}_{n,k}.$$

Reversing the above RH sum to the Mayer one by $e_{ij} \to 1 + f_{ij}$, we get

$$Z_n = \sum_k S^*_{n,k} I^{(\mathrm{M})}_{n,k},$$

where $S^*_{n,k}$ is the signed star content of the Mayer diagram $G_{n,k}$ (which is defined to be the same as that of the

corresponding RH diagram), e.g., $Z_4 = \square^{(M)} + \bowtie^{(M)} + \bowtie^{(M)} - 2\boxtimes^{(M)}$.

From the partition function, we can compute useful quantities as averages over biconnected configurations $\langle A\rangle_{bc} \equiv \int A(\mathbf{r}^n)\,\mathrm{bc}(\mathbf{r}^n)d\mathbf{r}^n/Z_n$. For example, from Equations (4) and (5), we get

$$B_n = \frac{1-n}{n!}\left\langle S^*(\mathbf{r})\right\rangle_{bc} Z_n. \tag{6}$$

2.4. Computation of the Ree–Hoover star content

In this section, we discuss several methods of computing the star content. The methods are equally applicable to the traditional MC integration method and the important sampling method developed later.

The first and direct method is well known [15] and we briefly review it. First, note that the star content $S(\mathbf{r}^n)$ depends on $\mathbf{r}^n$ only through the intrinsic RH diagram $G(\mathbf{r}^n)$. By the RH theorem, we only need to count the subgraphs of $G(\mathbf{r}^n)$ with even and odd numbers of missing f-bonds, and compute the difference. The method can be implemented in a depth-first search. That is, we recursively remove some set of f-bonds (but not vertices) in the subgraph until the subgraph is no longer biconnected; then, we add back some removed f-bonds and try to remove another set of f-bonds. During the search, we try to remove an f-bond $\{i, j\}$ only if both i and j have degrees greater than 2 because every vertex needs at least two adjacent vertices to make the subgraph without $\{i, j\}$ biconnected. The method works best for loose diagrams because its complexity grows exponentially with the number of f-bonds in the RH diagram.

An alternative method of evaluating the star content has recently been proposed [18]. Starting from the Boltzmann weight, the method collectively deducts the contributions from disconnected subgraphs and those from connected subgraphs with articulation points by two recursions [18]. This method depends weakly on the number of f-bonds, and works well for dense diagrams. It, however, requires memory of order 2^{n+1} (Appendix 1), so the highest affordable n is around 26 at present. In practice, we use this method only if the RH diagram has more than $2n-3$ f-bonds and if there is enough memory. While this method is very useful in low dimensions, the direct method is more often invoked in high dimensions, where configurations are mostly loose.

Both of the above methods can be time-consuming. For small n, one can avoid redundant calculations on recurring diagrams by using a pre-computed lookup table. Since an RH diagram of n vertices is fully specified by the $n \times n$ symmetric adjacency matrix, in which the element $\{i, j\}$ is 0 for an e-bond or 1 for an f-bond between vertices i and j, we can enumerate the $2^{n(n-1)/2}$ possibilities in advance [16,17]. This lookup table was used here for $n \le 8$. Higher order cases are more difficult. For $n = 9$ and 10, some of our preliminary simulations adopted a technique in Ref. [17], in which we trace the diagram by a short path of f-bonds and spare these f-bonds from the adjacency matrix in specifying the diagram (Appendix 2). Alternatively, by using the graph-isomorphism program NAUTY [27], we can recognise topologically equivalent diagrams, such as $\square$ and $\bowtie$, whose star contents are the same, and replace the lookup table by a smaller hash table [16]. The hash table was previously pre-computed, and the technique worked for n up to 10 [16]. For higher orders, we can exploit the fact that only a small fraction of geometrically permissible diagrams are frequently visited, and build the hash table on the fly to further reduce the memory usage (Appendix 3).

We can also avoid computing the star content by detecting if an RH diagram has a *clique separator*, which is a set of m ($m \le n-2$) pairwise-connected (in terms of f-bonds) vertices, whose removal along with the incident f-bonds disconnects the diagram. If an RH diagram has a clique separator, the star content is zero (Appendix 4). An $m = 1$ clique separator is an articulation point, and we call the $m = 2$ case a *separation f-bond*. Since f-bonds are readily enumerated, we try to detect separation f-bonds first, and, if it fails, then switch to the general algorithm by Tarjan *et al.* [28–31] (Appendix 5).

Finally, we list a few shortcuts to compute the star contents of some loose RH diagrams. If a biconnected diagram has exactly n f-bonds, it is a ring (for every vertex is of degree 2), and the star content is 1.0. If the diagram has $n+1$ f-bonds, there are two degree-3 vertices, and $n-2$ degree-2 vertices. The star content is 0, if the two degree-3 vertices are adjacent (and thus form a separation f-bond), or 1.0 otherwise. The pre-screening of the above loose diagrams is particularly useful in high dimensions.

2.5. Monte Carlo integration

The standard tool of evaluating virial coefficients is MC integration [5,14]. The idea is to treat the integral of an RH diagram as the probability of forming a compatible configuration. The probability is measured against that of generating a spanning tree of the RH diagram. A *spanning tree* is a connected subgraph of n vertices and $n-1$ f-bonds, e.g., a linear chain. To apply the method, configurations are first generated according to the spanning tree from the root to leaves. For example, in the case of a linear chain, we place one end of the chain at the origin, and randomly place each subsequent point within a unit sphere centred at the previous point. The intrinsic RH diagrams of the generated configurations, if biconnected, are then collected to estimate the respective integrals. The linear chain can cover many but not all diagrams for $n \ge 6$, and the number of spanning trees to cover all possible RH diagrams increases with n [16].

The MC integration method works well in low dimensions, especially for $n \le 10$. It, however, gets cumbersome as the dimensionality increases because the generated configurations become too loose to be biconnected [22], as

shown in Figure 1. An attractive alternative is the Mayer sampling strategy [23,24], which substitutes an importance sampling for the MC integrator. However, as an MC sampling only gives averaged quantities, which are ratios to the partition function, we need a reference value to get the absolute virial coefficient (as the partition function is usually unknown). The ring integral has been used as a convenient reference [24], and this approach is elaborated below.

2.6. Ring content

We define the *ring integral* R_n as the absolute value of the sum of all Mayer ring diagrams of n vertices:

$$R_n \equiv \frac{n-1}{n!}\left|\sum_{(v_1,v_2,\ldots,v_n)}\int f_{v_1v_2}f_{v_2v_3}\cdots f_{v_nv_1}d\mathbf{r}^n\right|, \quad (7)$$

where the summation is carried over all free circular permutations $(v_1, v_2, \ldots, v_n)$ of the vertices [here 'free' means direction-insensitive, e.g., (1, 2, 3, 4) is equivalent to (1, 4, 3, 2)]. Since all ring diagrams are isomorphic, the $n!/(2n)$ non-zero integrals all yield the same value, and $R_n = \frac{n-1}{2n}\int |f_{12}f_{23}\cdots f_{n1}|\, d\mathbf{r}^n$. We will use the integral as our reference, and its exact value is known [11,22]:

$$R_n/B_2^{n-1} = \Gamma(D/2)^{n-2}(1-\tfrac{1}{n})D^{n-1}2^{nD/2-D} \times \int_0^{+\infty}\left[J_{D/2}(k)/k^{D/2}\right]^n k^{D-1}dk, \quad (8)$$

where $J_\nu(z)$ and $\Gamma(z) = (z-1)!$ are the Bessel and gamma functions, respectively, and $B_2 = \sqrt{\pi}^D/[D\,\Gamma(D/2)]$ is the second virial coefficient. The integrals were evaluated using the program MATHEMATICA® [32].

By exchanging the order of the summation and integration in Equation (7), we get

$$R_n = -\frac{1-n}{n!}\int R(\mathbf{r}^n)\, d\mathbf{r}^n, \quad (9)$$

where

$$R(\mathbf{r}^n) \equiv \left|\sum_{(v_1,v_2,\ldots,v_n)} f_{v_1v_2}f_{v_2v_3}\cdots f_{v_nv_1}\right|, \quad (10)$$

which we call the *ring content*, gives the total contribution of the configuration $\mathbf{r}^n$ to the ring integral R_n. The domain of the integral in Equation (9) can be limited to the biconnected configurations since any diagram containing a ring subgraph is biconnected. So,

$$R_n = -\frac{1-n}{n!}\left\langle R(\mathbf{r}^n)\right\rangle_{\rm bc} Z_n. \quad (11)$$

From Equations (6) and (11), we have

$$\frac{B_n}{R_n} = \frac{\langle S^*(\mathbf{r}^n)\rangle_{\rm bc}}{-\langle R(\mathbf{r}^n)\rangle_{\rm bc}}, \quad (12)$$

which allows us to compute the virial coefficient B_n from the averaged star and ring contents [24].

We can use Equation (10) to compute the ring content by counting the number of Hamiltonian cycles, or free circular permutations $(v_1, v_2, \ldots, v_n)$, such that $\{v_1, v_2\}$, …, $\{v_{n-1}, v_n\}$, and $\{v_n, v_1\}$ are all f-bonds in the intrinsic RH diagram $G(\mathbf{r}^n)$. The counting is readily implemented as a depth-first search over successive f-bonds $\{v_k, v_{k+1}\}$, with k being the depth. While the algorithm works well for loose diagrams, its complexity can go up to $O(n!)$ in the worst case. Some alternative algorithms for dense diagrams are discussed in Appendix 6. We can pre-screen a few loose configurations as before. If there are exactly n f-bonds, the ring content is 1.0. If there are exactly $n+1$ f-bonds, the ring content is 1.0 if the two degree-3 vertices are adjacent, or 0 otherwise.

Finally, we give another interpretation of Equation (10). We can expand all Mayer *ring* diagrams (instead of all Mayer diagrams) in terms of RH diagrams. Then, the RH theorem states that the ring content, which is the absolute coefficient before an RH diagram G, is equal to the number of ring subgraphs of G, agreeing with Equation (10). Practically, this means that the ring content can be computed along with the star content whenever the direct method discussed in Section 2.4 is used.

2.7. Metropolis algorithm

We now give a Metropolis algorithm, called Algorithm R, based on Equation (12).

(1) Prepare an n -point biconnected configuration in D dimensions, e.g., a planar ring of n points. Repeat steps (2)–(4) in each MC step.
(2) Displace each coordinate of a random point i by a uniform random number in $[-a, a]$, with $a = 1.5/D$.
(3) If the new configuration is biconnected, accept it. Otherwise, return to the unperturbed configuration.
(4) Compute and save the star and ring contents of the current configuration after a period of equilibration.
(5) Compute the virial coefficient B_n from Equation (12).

The above algorithm is a variant of one previously published [24], and we discuss some changes below. The current algorithm uniformly samples all biconnected configurations, while the original algorithm, which is designed for soft potentials, would use the absolute star contents as the weight, and hence ignore configurations $\mathbf{r}^n$ with zero star content in $G(\mathbf{r}^n)$ [24]. Our modification is necessary because, in a hard-sphere fluid, a configuration $\mathbf{r}^n$ with zero star content may have non-zero ring content in $G(\mathbf{r}^n)$. Thus, these configurations have to be included in the sampling as well [24]. In Appendix 7, we show that

the modified sampling is ergodic even if only single-point trial moves are used. Thus, the multiple-point trial moves [24] are unnecessary here. Second, the magnitude a of displacement in step (2) is chosen to make the acceptance ratio of the configuration trials close to 50%, as recommended previously [24]. Further, the sampling frequency in step (4) is adjustable: calculating the star and ring contents in every MC step is neither necessary for there are correlations among successive steps, nor efficient if the calculation is expensive. In practice, we usually computed the star and ring contents (cf. Sections 2.4 and 2.6) in every step for low-order ($n < 14$) or low-dimensional ($D < 6$) cases, or at a 10-step interval otherwise.

The above algorithm uses the ring integral as the reference integral and thus requires one to compute the ring content. For low-dimensional and high-order coefficients, which involve many dense diagrams, it is more efficient to change the reference integral to the partition function, which can be obtained from the grand ensemble technique described below. The change not only saves the time of computing the ring content, but also allows a larger dynamic hash table (Appendix 3) to evaluate the star content more efficiently. In this variant, or Algorithm Z below, we compute and save only the star content in step (4), and replace, in step (5), Equation (12) by Equation (6).

2.8. *Grand ensemble sampling*

In this and the next sections, we develop a grand ensemble technique that is able to compute the virial coefficients for a range of orders n within a single MC simulation. As in the standard grand-canonical ensemble technique [33], n is treated as a variable (but we drop 'canonical' for convenience). Here, however, we sample a flat distribution along n instead of the one under a fixed 'pressure'. Despite the conceptual simplicity, our final algorithm is somewhat involved, owing to the use of a heat-bath technique [34] and the target free-energy technique [35] for better performance at large n and D. To set the solution in context, we first introduce in this section two simpler variants before the full version in the next section.

First, recall that once the partition function Z_n is known, we can compute the virial coefficient B_n from the average $\langle S^*(\mathbf{r})\rangle_{\rm bc}$ by Equation (6) without resorting to the ring integral. In the grand ensemble, we can compute the ratios Z_{k+1}/Z_k from the mutual transition rates between the configurations of k and $k+1$ points by, e.g., the Bennett acceptance ratio method [36]. The partition function is given by

$$Z_n = Z_{n_0} \prod_{k=n_0}^{n-1} Z_{k+1}/Z_k, \tag{13}$$

where $n_0 = 3$ for $D > 12$ or 4 otherwise, because $Z_3 = 3R_3$ is always known [3], and Z_4 is known for $D \le 12$ [7,11–13] (cf. Supplemental material for explicit values). The ratio Z_{k+1}/Z_k obtained here is also useful in determining the radius of convergence of the virial series (cf. Section 2.10). Another way to obtain Z_n is through Equation (11).

Although we can directly compute the virial coefficients from a grand ensemble simulation, we often run such a simulation just to get the partition function up to a certain order $n_{\max}$ in low dimensions. In this case, we do not compute the star or ring contents. The obtained partition function is then used in specialised and independent simulations of Algorithm Z for individual orders $n < n_{\max}$. This strategy, however, is necessary only for high-order cases, where the partition function calculation converges much faster than the average of the signed star content, whose convergence is hampered by heavy cancellation among dense diagrams.

We now describe a simplest version of the algorithm.

Algorithm G1: Given a range of orders $[n_{\min}, n_{\max}]$, we prepare a biconnected configuration of $n = n_{\min}$ points. In each MC step, we randomly choose whether to perform a configuration sampling step, as described in Section 2.7, or to change n. In the latter case, we decide randomly whether to increase or to decrease n by 1. We immediately reject an attempt to move n outside of $[n_{\min}, n_{\max}]$. To increase n, we pick a random point i in the configuration, and place the new point $n+1$ randomly in a D-dimensional sphere of radius R centred at point i (the parameter R is to be determined). We accept the enlarged $n+1$-point configuration if it is biconnected, or equivalently (Appendix 8), if there are two points k, l ($k, l \le n$) that satisfy $r_{k,n+1} < 1$ and $r_{l,n+1} < 1$. To shrink a configuration of $n+1$ points, we randomly pick two points i and j, and remove point i if $r_{i,j} < R$ and if the configuration without i is biconnected.

It is straightforward to check that the algorithm obeys detailed balance [33,34], and the acceptance ratios satisfy (Appendix 9)

$$\frac{\langle p_{n\to n+1}\rangle}{\langle p_{n+1\to n}\rangle} = \frac{Z_{n+1}/(Z_n V_D)}{R^D}, \tag{14}$$

where $V_D = \sqrt{\pi}^D/\Gamma(D/2+1)$. This gives an estimate for the *inverse activity* $Z_{n+1}/(Z_n V_D)$. By detailed balance, the ratio in Equation (14) equals the population ratio P_{n+1}/P_n in the two ensembles, which should be unity ideally. So, the (n-dependent) optimal value of R is $R = \sqrt[D]{Z_{n+1}/(Z_n V_D)}$. To reach this value, we periodically update R in a short preparation run as

$$R^{\rm new} = R^{\rm old}\left(\langle p_{n\to n+1}\rangle/\langle p_{n+1\to n}\rangle\right)^{1/D}. \tag{15}$$

A drawback of the above algorithm is that the acceptance ratio $\langle p_{n+1\to n}\rangle$ quickly deteriorates with increasing n. This is because during the point-removal trial, two random vertices i and j are unlikely to be separated by a distance less than R, which must be small enough to maintain a

reasonable $\langle p_{n\to n+1}\rangle$ for the point-addition trial. We can, however, improve the algorithm by choosing the pair $\{i, j\}$ more carefully using the following heat-bath technique.

Algorithm G2: We will only modify the procedures of adding and removing points, and the rest are the same as Algorithm G1. To remove a point from an $n+1$-point configuration, we first list all *ordered* pairs (i, j) that (1) satisfy $r_{i,j} < R$ and (2) leave the configuration biconnected once point i is removed. If the list is not empty and contains $N_p(\mathbf{r}^{n+1})$ pairs, we randomly pick an ordered pair from the list, and accept it by $\min\{1, N_p(\mathbf{r}^{n+1})/\bar{N}_p\}$, where the parameter $\bar{N}_p$ is to be determined. To enlarge an n-point configuration, we attach a new point around a random existing point i within the uniform sphere of radius R as before. However, now even if the new configuration is biconnected, we only accept it with probability $\min\{1, \bar{N}_p/N_p(\mathbf{r}^{n+1})\}$, where $N_p(\mathbf{r}^{n+1})$ is the number of ordered pairs that satisfy the above conditions (1) and (2) in the enlarged configuration $\mathbf{r}^{n+1}$.

Again, detailed balance can be straightforwardly checked, and the acceptance ratios satisfy (Appendix 9)

$$\frac{\langle p_{n\to n+1}\rangle}{\langle p_{n+1\to n}\rangle} = \frac{\bar{N}_p Z_{n+1}/(Z_n V_D)}{n(n+1)R^D}. \tag{16}$$

The parameters are optimised during the preparation run as follows. We set $\bar{N}_p$ as the average of non-zero values of $N_p(\mathbf{r}^{n+1})$ in the point-removal move to make the probability $\min\{1, N_p(\mathbf{r}^{n+1})/\bar{N}_p\}$ close to 1. The optimal radius R, which makes the ratio in Equation (16) unity and hence achieves a flat distribution across n, can still be obtained from Equation (15).

Although Algorithm G2 works well for large n with a moderate dimension D, it is inefficient in higher dimensions in bridging the geometric difference between configurations of successive orders. We solve this problem in the full version of the algorithm (Section 2.9) by introducing a few intermediate states and actively pulling the configuration along a reaction coordinate of n.

2.9. *Grand ensemble sampling with intermediate ensembles*

We now describe the full grand ensemble algorithm, Algorithm G3. First, to facilitate the transition between the n-point and $n+1$-point configurations, we use the distance R between a non-separation pair in the $n+1$-point configuration as the reaction coordinate. We will set up M intermediate ensembles between the n-point ensemble and the $n+1$-point ensemble, with $R^1 < \cdots < R^M$ and $R^M = 1.0$. We label an intermediate ensemble by the superscript m, and $m = 0$ and $M+1$ mean the regular n-point and $n+1$-point ensembles, respectively. The partition function of the mth intermediate ensemble Z^m is

$$Z^m = \int \Theta(R^m - \mathbf{r}_{i,j})c(\mathbf{r}^{n+1}_{-\{i,j\}})\mathrm{bc}(\mathbf{r}^{n+1})d\mathbf{r}^{n+1} \qquad (1 \le m \le M),$$

where $\mathbf{r}^{n+1}_{-S}$ means the configuration formed by excluding all points in the set S (in this case points i and j) from the $n+1$-point configuration $\mathbf{r}^{n+1}$, and $c(\mathbf{r}^{n+1}_{-\{i,j\}})$ is 1.0 if $\mathbf{r}^{n+1}_{-\{i,j\}}$ is connected, i.e., $\{i, j\}$ is not a separation pair in $G(\mathbf{r}^{n+1})$, or 0 otherwise.

The nature of the reaction coordinate can be seen from two limits. In the limit of $R \to 0$, point i collapses on point j: $\mathbf{r}_i \to \mathbf{r}_j$. Then, after the removal of point i, the n-point configuration $\mathbf{r}^{n+1}_{-\{i\}}$ is biconnected. This is because the biconnectivity $\mathrm{bc}(\mathbf{r}^{n+1})$ requires the intrinsic RH diagram $G(\mathbf{r}^{n+1}_{-\{v\}})$ $(v \neq i, j)$, hence $G(\mathbf{r}^{n+1}_{-\{i,v\}})$, to be connected, and $c(\mathbf{r}^{n+1}_{-\{i,j\}})$ ensures that $G(\mathbf{r}^{n+1}_{-\{i,j\}})$ is connected. In the other limit of $R \to +\infty$, the distance constraint $\Theta(R - r_{i,j})$ is trivially satisfied, and a biconnected configuration $\mathbf{r}^{n+1}$ in the regular ensemble can reach the intermediate ensemble as long as there is a non-separation pair $\{i, j\}$. Thus, we can switch from a biconnected configuration of n points to that of $n+1$ points by increasing R from zero. It is, however, unnecessary to reach the two limits in a practical setting.

Sampling in an intermediate ensemble is similar to that in a regular ensemble, but the additional constraints $\Theta(R - r_{i,j})$ and $c(\mathbf{r}^{n+1}_{-\{i,j\}})$ need to be checked in the trial moves. The star and ring contents are not computed if the current ensemble is an intermediate one. The rules of transitions among intermediate and regular ensembles are specified below.

(I) A transition from the n-point regular ensemble to the $n+1$-point intermediate ensemble $(m : 0 \to 1)$ is similar to the $n \to n+1$ step in Algorithm G1. We place the new point $n+1$ randomly in a sphere of radius R^1 centred at a random point i. If the enlarged configuration $\mathbf{r}^{n+1}$ is biconnected, we accept it with probability $\min\{1, 1/\zeta\}$. The parameter ζ is to be determined later. Note, since i is not an articulation point of $\mathbf{r}^n$, $\{i, n+1\}$ is not a separation pair of $\mathbf{r}^{n+1}$, and the constraint $c(\mathbf{r}^{n+1}_{-\{i,n+1\}})$ is satisfied. The reverse transition $(m : 1 \to 0)$ is executed as follows. Let $\{i, j\}$ be the pair under the distance constraint $\Theta(R - r_{ij})$ in the intermediate ensemble of $n+1$ points. If the n-point configuration $\mathbf{r}^{n+1}_{-\{i\}}$ is biconnected, we remove point i with probability $\min\{1, \zeta\}$. We choose i and j equally by symmetry.

(II) From the $n+1$-point regular ensemble to the last $n+1$-point intermediate ensemble $(m : M+1 \to M)$. Since $R^M = 1$, the latter ensemble requires a non-separation f-bond to satisfy $\Theta(1 - r_{i,j})c(\mathbf{r}^{n+1}_{-\{i,j\}}) = 1$. This can be executed efficiently

using the heat-bath technique introduced in Algorithm G2. We randomly pick an f-bond, and if it is not a separation pair, we accept the transition with probability $\min\{1, N_f(\mathbf{r}^{n+1})/\overline{N}_f\}$, where $N_f(\mathbf{r}^{n+1})$ is the number of f-bonds in the intrinsic RH diagram $G(\mathbf{r}^{n+1})$. For simplicity, we set the parameter $\overline{N}_f = n+1$, instead of the average of $N_f(\mathbf{r}^{n+1})$ (the value $n+1$ is equal to the number of f-bonds in the $n+1$-point ring). For the reverse transition ($m : M \to M+1$), we simply count the number of f-bonds $N_f(\mathbf{r}^{n+1})$ in the intrinsic RH diagram $G(\mathbf{r}^{n+1})$, and accept the move with probability $\min\{1, \overline{N}_f/N_f(\mathbf{r}^{n+1})\}$.

(III) For two, say, the mth and m'th ($1 \le m,\ m' \le M$) intermediate ensembles of $n+1$ points, we use a targeted free-energy technique [35] to transform the configuration $\mathbf{r}^n$ along the reaction coordinate before making a transition. We first displace point i in the distance-constrained pair $\{i, j\}$ as

$$\begin{cases} \mathbf{r}'_i = \mathbf{r}_j + s(\mathbf{r}_i - \mathbf{r}_j), \\ \mathbf{r}'_k = \mathbf{r}_k \quad \text{for } k \neq i, \end{cases} \tag{17}$$

and then accept the configuration $\mathbf{r}'^{n+1}$ if $|\mathbf{r}'_i - \mathbf{r}'_j| < R^{m'}$ and if the configuration $\mathbf{r}'^{n+1}$ is biconnected. The parameter s is to be determined later. For the reverse transition $m' \to m$, we follow the same protocol but use the inverse scaling factor $s' = 1/s$. For simplicity, we only allow transitions between neighbouring ensembles, i.e., $m' = m \pm 1$. So, it suffices to specify $M-1$ values of s^m for all pairs $m \leftrightarrow m+1$, with $m = 1, \ldots, M-1$.

The high acceptance ratios (which are close to 1.0) in type II transitions allow us to implement the last ($m = M$) intermediate ensemble as a 'virtual' ensemble, which can be skipped over during the transitions along m. A transition $M+1 \to M$ is immediately followed by a transition $M \to M-1$, and if either transition is rejected, so is the composite transition. The two reverse transitions $M-1 \to M$ and $M \to M+1$ are similarly combined.

The acceptance ratios of types I, II, and III transitions are (Appendix 9)

$$\frac{\langle p^{0\to1}\rangle}{\langle p^{1\to0}\rangle} = \frac{Z^1/(Z^0 V_D)}{\zeta (R^1)^D}, \tag{18}$$

$$\frac{\langle p^{M\to M+1}\rangle}{\langle p^{M+1\to M}\rangle} = \frac{\overline{N}_f}{n(n+1)/2}\frac{Z^{M+1}}{Z^M}, \tag{19}$$

and

$$\frac{\langle p^{m\to m+1}\rangle}{\langle p^{m+1\to m}\rangle} = \frac{Z^{m+1}}{(s^m)^D Z^m}, \quad m = 1, \ldots, M-1. \tag{20}$$

From Equations (18), (19), and (20), we get

$$\prod_{m=0}^{M} \frac{\langle p^{m\to m+1}\rangle}{\langle p^{m+1\to m}\rangle} = \frac{2\overline{N}_f}{n(n+1)\zeta}\frac{Z^{M+1}/(Z^0 V_D)}{(R^1)^D \prod_{m=1}^{M-1}(s^m)^D}, \tag{21}$$

which gives an estimate of the inverse activity $Z^{M+1}/(Z^0 V_D) = Z_{n+1}/(Z_n V_D)$.

The parameters are updated periodically in a preliminary preparation run to achieve the ideal values. The parameter ζ is updated according to

$$\zeta_{\text{new}} = \zeta_{\text{old}}\langle p^{0\to1}\rangle/\langle p^{1\to0}\rangle,$$

to balance the populations of the two ensembles at $m = 0$ and 1. The parameter s^m is updated according to

$$s^m_{\text{new}} = s^m_{\text{old}}\left(\langle p^{m\to m+\delta}\rangle/\langle p^{m+\delta\to m}\rangle\right)^{1/D},$$

where $\delta = 1$ if $1 \le m < M-1$, or 2 if $m = M-1$, as the $m = M$ ensemble is skipped over. After convergence, we can achieve a relatively flat distribution among all non-virtual ensembles. The parameters R^m ($1 \le m \le M-1$) are updated more carefully. If the acceptance ratios $\langle p^{m\to m-1}\rangle$ and $\langle p^{m-1\to m}\rangle$ differ by less than 10%, and the same holds for $\langle p^{m+\delta\to m}\rangle$ and $\langle p^{m\to m+\delta}\rangle$, we apply $R^m_{\text{new}} = R^m_{\text{old}}(\bar{p}^{m-1,m}/\bar{p}^{m,m+\delta})^{4/D}$, where $\bar{p}^{m,m'} \equiv \left(\langle p^{m\to m'}\rangle + \langle p^{m'\to m}\rangle\right)/2$, and the exponent $4/D$ is arbitrary. Initially, we set $R^m = m/M$ and $s^m = R^{m+1}/R^m$.

2.10. *Radius of convergence*

We can estimate the radius of convergence $\rho = \lim_{n\to\infty} |B_n/B_{n+1}|$ of the virial series following two routes. The first is the ring route. By Equation (12), we have

$$\rho^{(R)} V_D = \lim_{n\to\infty} \frac{q_n}{q_{n+1}}\frac{2R_n/B_2^{n-1}}{R_{n+1}/B_2^n} = \lim_{n\to\infty}\frac{q_n}{q_{n+1}}, \tag{22}$$

where $V_D = 2B_2$, and $q_n \equiv |\langle S^*(\mathbf{r}^n)\rangle_{\text{bc}}|/\langle R(\mathbf{r}^n)\rangle_{\text{bc}}$. The second step follows from the asymptotic limit of Equation (8) (cf. Appendix 10). We then fit q_n against $-c_1 n - c_2 - c_3/n$ by regression, and $\rho^{(R)} V_D = \exp(c_1)$.

The second is the activity or partition function route. By Equation (6), we have

$$\rho^{(Z)} V_D = \lim_{n\to\infty}\left[\frac{n(V_D Z_n)}{Z_{n+1}}\right]\left|\frac{\langle S^*(\mathbf{r}^n)\rangle_{\text{bc}}}{\langle S^*(\mathbf{r}^{n+1})\rangle_{\text{bc}}}\right|. \tag{23}$$

Here, $Z_{n+1}/(V_D Z_n)$ can be estimated from Equation (21) in the grand ensemble. We fit the two factors on the right side of Equation (23) as $Z_{n+1}/(V_D Z_n) \approx (a_1 n + a_2)/(1 + a_3/n)$ and $\log|\langle S^*\rangle| \approx -b_1 n - b_2 - b_3/n$. Then, $\rho^{(Z)} V_D = \exp(b_1)/a_1$. Note that

the orders of the above rational functions are chosen out of convenience, and thus adjustable.

3. Results

3.1. Virial coefficients

We first applied Algorithms R and Z discussed in Section 2.7 to compute a few virial coefficients of hard-sphere fluids in dimensions $D \leq 8$. In simulations of Algorithm Z, the values of the partition functions were pre-computed to relative errors less than 0.01%. The results are shown in Table 1 (cf. Supplemental data for additional details). For $D \geq 4$, we improved the precision of all existing hard-sphere virial coefficients from the literature and obtained a few new results for orders $n > 10$. The new values agreed well with the previous extrapolated values [16,37,38]. The improvement in either precision or attained order was more significant in

Table 1. Virial coefficients in low dimensions.

D_n	B_n/B_2^{n-1}	D_n	B_n/B_2^{n-1}	D_n	B_n/B_2^{n-1}	D_n	B_n/B_2^{n-1}
2_5	0.33355619(75) 0.33355604(4)[§]	2_6	0.19884472(67) 0.19884456(34)[#]	2_7	0.11487473(96) 0.11487625(44)[#]	2_8	0.0649916(23) 0.0649896(5)[#]
2_9	0.0362229(29) 0.0362202(11)[#]	2_{10}	0.0199521(62)[¶] 0.0199537(80)[*]	2_{11}	0.010933(19)[¶] 0.01089[†]	2_{12}	$5.862(53) \times 10^{-3}$[¶] 5.90×10^{-3}[†]
2_{13}	$3.16(16) \times 10^{-3}$[¶] 3.18×10^{-3}[†]	2_{14}	$1.82(51) \times 10^{-3}$[¶] 1.70×10^{-3}[†]				
3_5	0.11025161(13) 0.1102518(4)[#]	3_6	0.03888188(23) 0.03888232(35)[#]	3_7	0.01302331(28) 0.0130228(4)[#]	3_8	$4.18300(52) \times 10^{-3}$ $4.1833(6) \times 10^{-3}$[#]
3_9	$1.3093(11) \times 10^{-3}$ $1.3092(12) \times 10^{-3}$[#] $1.3094(13) \times 10^{-3}$[*]	3_{10}	$4.047(14) \times 10^{-4}$[¶] $4.035(15) \times 10^{-4}$[*] $4.05(4) \times 10^{-4}$[#]	3_{11}	$1.224(35) \times 10^{-4}$[¶] $1.22(4) \times 10^{-4}$[¤] 1.22×10^{-4}[†] $1.24(1) \times 10^{-4}$[‡] $1.25(1) \times 10^{-4}$[‖]	3_{12}	$4.0(9) \times 10^{-5}$[¶] $2.6(7) \times 10^{-5}$[¤] 3.64×10^{-5}[†] $3.76(9) \times 10^{-5}$[‡] $3.74(7) \times 10^{-5}$[‖]
4_5	0.03570438(12) 0.0357046(3)[#]	4_6	$7.73280(16) \times 10^{-3}$ $7.7331(4) \times 10^{-3}$[#]	4_7	$1.43077(20) \times 10^{-3}$ $1.4309(5) \times 10^{-3}$[#]	4_8	$2.9051(29) \times 10^{-4}$ $2.900(8) \times 10^{-4}$[#]
4_9	$4.57(6) \times 10^{-5}$ $4.45(11) \times 10^{-5}$[#]	4_{10}	$1.06(8) \times 10^{-5}$[¶] $1.13(31) \times 10^{-5}$[*]				
5_5	$1.295219(16) \times 10^{-2}$ $1.29551(13) \times 10^{-2}$[*]	5_6	$9.8184(19) \times 10^{-4}$ $9.815(14) \times 10^{-4}$[*]	5_7	$4.1645(21) \times 10^{-4}$ $4.162(19) \times 10^{-4}$[*]	5_8	$-1.1267(43) \times 10^{-4}$ $-1.120(20) \times 10^{-4}$[*]
5_9	$7.886(45) \times 10^{-5}$ $7.47(26) \times 10^{-5}$[*]	5_{10}	$-4.68(10) \times 10^{-5}$[¶] $-4.92(48) \times 10^{-5}$[*]	5_{11}	$3.09(11) \times 10^{-5}$[¶] 3×10^{-5}[†]	5_{12}	$-2.29(23) \times 10^{-5}$[¶] -4×10^{-5}[†]
6_5	$7.52384(18) \times 10^{-3}$ $7.5231(11) \times 10^{-3}$[*]	6_6	$-1.74015(21) \times 10^{-3}$ $-1.7385(13) \times 10^{-3}$[*]	6_7	$1.30523(30) \times 10^{-3}$ $1.3066(18) \times 10^{-3}$[*]	6_8	$-8.9254(27) \times 10^{-4}$ $-8.950(30) \times 10^{-4}$[*]
6_9	$6.6920(38) \times 10^{-4}$ $6.673(45) \times 10^{-4}$[*]	6_{10}	$-5.294(10) \times 10^{-4}$[¶] $-5.25(16) \times 10^{-4}$[*]	6_{11}	$4.382(21) \times 10^{-4}$[¶] 4.32×10^{-4}[†]	6_{12}	$-3.766(43) \times 10^{-4}$[¶] -3.68×10^{-4}[†]
6_{13}	$3.19(13) \times 10^{-4}$[¶] 3.2×10^{-4}[†]	6_{14}	$-2.97(17) \times 10^{-4}$[¶] -2.9×10^{-4}[†]	6_{15}	$2.67(47) \times 10^{-4}$[¶] 2.7×10^{-4}[†]	6_{16}	$-2.6(7) \times 10^{-4}$[¶] -2.5×10^{-4}[†]
7_5	$7.07178(15) \times 10^{-3}$ $7.0724(10) \times 10^{-3}$[*]	7_6	$-3.51139(16) \times 10^{-3}$ $-3.5121(11) \times 10^{-3}$[*]	7_7	$2.53868(17) \times 10^{-3}$ $2.5386(16) \times 10^{-3}$[*]	7_8	$-1.99410(32) \times 10^{-3}$ $-1.9937(28) \times 10^{-3}$[*]
7_9	$1.69032(31) \times 10^{-3}$ $1.6869(41) \times 10^{-3}$[*]	7_{10}	$-1.5178(5) \times 10^{-3}$ $-1.514(14) \times 10^{-3}$[*]	7_{11}	$1.4261(8) \times 10^{-3}$[¶] 1.43×10^{-3}[†]	7_{12}	$-1.3877(9) \times 10^{-3}$[¶] -1.40×10^{-3}[†]
7_{13}	$1.3888(16) \times 10^{-3}$[¶] 1.42×10^{-3}[†]	7_{14}	$-1.4363(36) \times 10^{-3}$[¶] -1.45×10^{-3}[†]	7_{15}	$1.508(8) \times 10^{-3}$[¶] 1.5×10^{-3}[†]	7_{16}	$-1.631(15) \times 10^{-3}$[¶] -1.6×10^{-3}[†]
7_{17}	$1.84(5) \times 10^{-3}$[¶] 1.8×10^{-3}[†]	7_{18}	$-1.94(12) \times 10^{-3}$[¶] -2.1×10^{-3}[†]	7_{19}	$2.57(23) \times 10^{-3}$[¶]	7_{20}	$-2.29(34) \times 10^{-3}$[¶]
8_5	$7.43187(14) \times 10^{-3}$ $7.43092(93) \times 10^{-3}$[*]	8_6	$-4.51452(7) \times 10^{-3}$ $-4.5164(11) \times 10^{-3}$[*]	8_7	$3.41415(7) \times 10^{-3}$ $3.4149(15) \times 10^{-3}$[*]	8_8	$-2.86716(11) \times 10^{-3}$ $-2.8624(26) \times 10^{-3}$[*]
8_9	$2.60397(16) \times 10^{-3}$ $2.5969(38) \times 10^{-3}$[*]	8_{10}	$-2.51041(38) \times 10^{-3}$ $-2.511(13) \times 10^{-3}$[*]	8_{11}	$2.5367(6) \times 10^{-3}$ 2.56×10^{-3}[†]	8_{12}	$-2.6592(10) \times 10^{-3}$ -2.72×10^{-3}[†]
8_{13}	$2.8773(13) \times 10^{-3}$[¶] 2.97×10^{-3}[†]	8_{14}	$-3.1924(21) \times 10^{-3}$[¶] -3.4×10^{-3}[†]	8_{15}	$3.6299(42) \times 10^{-3}$[¶] 4.0×10^{-3}[†]	8_{16}	$-4.202(8) \times 10^{-3}$[¶] -4.4×10^{-3}[†]
8_{17}	$4.950(11) \times 10^{-3}$[¶] 5.5×10^{-3}[†]	8_{18}	$-5.896(21) \times 10^{-3}$[¶] -6×10^{-3}[†]	8_{19}	$7.18(12) \times 10^{-3}$[¶]	8_{20}	$-8.72(23) \times 10^{-3}$[¶]
8_{21}	$1.028(49) \times 10^{-2}$[¶]	8_{22}	$-1.15(9) \times 10^{-2}$[¶]	8_{23}	$1.56(18) \times 10^{-2}$[¶]	8_{24}	$-2.02(45) \times 10^{-2}$[¶]

Notes: The superscript '[¶]' indicates the values obtained partially or entirely from simulations via Algorithm Z; other values were exclusively obtained from simulations via Algorithm R. The best literature values from Refs. [10][§], [17][#], [16][*], and [18][¤] are also shown. The values predicted from interpolating lower order virial coefficients are indicated by the superscripts '[†]' [16], '[‡]' [37], and '[‖]' [38].

Table 2. First eight virial coefficients in high dimensions.

D	B_4/B_2^3	B_5/B_2^4	B_6/B_2^5	B_7/B_2^6	B_8/B_2^7	$M^\dagger$	$T^\ddagger$
9	$-8.58057(40) \times 10^{-3}$	$7.44248(34) \times 10^{-3}$	$-4.79144(26) \times 10^{-3}$	$3.71459(37) \times 10^{-3}$	$-3.22289(44) \times 10^{-3}$	0	111
	$-8.580798\ldots \times 10^{-3\#}$	$7.4389(25) \times 10^{-3\square}$	$-4.7936(29) \times 10^{-3\square}$	$3.7159(42) \times 10^{-3\square}$	$-3.2223(56) \times 10^{-3\square}$		
10	$-0.01096172(26)$	$6.97418(26) \times 10^{-3}$	$-4.52937(25) \times 10^{-3}$	$3.52830(25) \times 10^{-3}$	$-3.09275(34) \times 10^{-3}$	0	112
	$-0.0109624743\ldots{}^*$	$6.969(5) \times 10^{-3\S}$	$-4.52(1) \times 10^{-3\S}$				
11	$-0.01133700(23)$	$6.18156(12) \times 10^{-3}$	$-3.95830(17) \times 10^{-3}$	$3.05131(13) \times 10^{-3}$	$-2.66039(15) \times 10^{-3}$	0	111
	$-0.0113371985\ldots{}^\#$	$6.176(4) \times 10^{-3\S}$	$-3.95(1) \times 10^{-3\S}$				
12	$-0.01067017(20)$	$5.24835(13) \times 10^{-3}$	$-3.26859(9) \times 10^{-3}$	$2.46516(9) \times 10^{-3}$	$-2.11265(14) \times 10^{-3}$	0	108
	$-0.01067028055\ldots{}^*$	$5.224(4) \times 10^{-3\S}$	$-3.261(7) \times 10^{-3\S}$				
13	$-9.52485(13) \times 10^{-3}$	$4.31120(9) \times 10^{-3}$	$-2.58675(7) \times 10^{-3}$	$1.89160(8) \times 10^{-3}$	$-1.57947(11) \times 10^{-3}$	0	109
	$-9.523(2) \times 10^{-3\S}$	$4.307(3) \times 10^{-3\S}$	$-2.580(6) \times 10^{-3\S}$				
14	$-8.22123(13) \times 10^{-3}$	$3.45105(8) \times 10^{-3}$	$-1.98077(7) \times 10^{-3}$	$1.394424(43) \times 10^{-3}$	$-1.12638(5) \times 10^{-3}$	0	108
	$-8.220(2) \times 10^{-3\S}$	$3.448(3) \times 10^{-3\S}$	$-1.975(4) \times 10^{-3\S}$				
15	$-6.93430(12) \times 10^{-3}$	$2.70711(5) \times 10^{-3}$	$-1.47762(4) \times 10^{-3}$	$9.95952(31) \times 10^{-4}$	$-7.73840(31) \times 10^{-4}$	1	57
	$-6.934(2) \times 10^{-3\S}$	$2.705(2) \times 10^{-3\S}$	$-1.472(3) \times 10^{-3\S}$				
16	$-5.75247(6) \times 10^{-3}$	$2.088671(35) \times 10^{-3}$	$-1.07930(2) \times 10^{-3}$	$6.93172(14) \times 10^{-4}$	$-5.15771(18) \times 10^{-4}$	1	65
17	$-4.71236(6) \times 10^{-3}$	$1.590203(23) \times 10^{-3}$	$-7.74830(16) \times 10^{-4}$	$4.72308(15) \times 10^{-4}$	$-3.35326(18) \times 10^{-4}$	1	63
18	$-3.82300(3) \times 10^{-3}$	$1.197257(25) \times 10^{-3}$	$-5.48250(12) \times 10^{-4}$	$3.16195(8) \times 10^{-4}$	$-2.13544(7) \times 10^{-4}$	1	62
19	$-3.07760(4) \times 10^{-3}$	$8.93015(16) \times 10^{-4}$	$-3.83262(5) \times 10^{-4}$	$2.085811(32) \times 10^{-4}$	$-1.33663(3) \times 10^{-4}$	1	57
20	$-2.46217(3) \times 10^{-3}$	$6.60869(9) \times 10^{-4}$	$-2.65182(6) \times 10^{-4}$	$1.358662(19) \times 10^{-4}$	$-8.24429(20) \times 10^{-5}$	1	60
	$-2.4621(7) \times 10^{-3\S}$	$6.605(7) \times 10^{-4\S}$	$-2.632(7) \times 10^{-4\S}$				
25	$-7.57933(7) \times 10^{-4}$	$1.347663(24) \times 10^{-4}$	$-3.76242(7) \times 10^{-5}$	$1.392319(21) \times 10^{-5}$	$-6.30009(14) \times 10^{-6}$	2	35
	$-7.580(3) \times 10^{-4\S}$	$1.348(2) \times 10^{-4\S}$	$-3.72(1) \times 10^{-5\S}$				
30	$-2.19569(2) \times 10^{-4}$	$2.514370(26) \times 10^{-5}$	$-4.75209(6) \times 10^{-6}$	$1.241925(13) \times 10^{-6}$	$-4.11652(4) \times 10^{-7}$	2	32
	$-2.196(1) \times 10^{-4\S}$	$2.515(6) \times 10^{-5\S}$	$-4.69(3) \times 10^{-6\S}$				
35	$-6.16116(6) \times 10^{-5}$	$4.46225(5) \times 10^{-6}$	$-5.62806(5) \times 10^{-7}$	$1.028674(17) \times 10^{-7}$	$-2.48279(3) \times 10^{-8}$	2	30
	$-6.162(3) \times 10^{-5\S}$	$4.47(1) \times 10^{-6\S}$	$-5.55(5) \times 10^{-7\S}$				
40	$-1.69682(1) \times 10^{-5}$	$7.68159(10) \times 10^{-7}$	$-6.41812(6) \times 10^{-8}$	$8.17545(7) \times 10^{-9}$	$-1.43527(1) \times 10^{-9}$	3	21
	$-1.697(1) \times 10^{-5\S}$	$7.69(3) \times 10^{-7\S}$	$-6.30(9) \times 10^{-8\S}$				
45	$-4.61835(5) \times 10^{-6}$	$1.297053(14) \times 10^{-7}$	$-7.15360(5) \times 10^{-9}$	$6.34654(4) \times 10^{-10}$	$-8.11252(5) \times 10^{-11}$	3	22
	$-4.618(4) \times 10^{-6\S}$	$1.298(7) \times 10^{-7\S}$	$-7.0(2) \times 10^{-9\S}$				
50	$-1.24719(1) \times 10^{-6}$	$2.162450(15) \times 10^{-8}$	$-7.86216(5) \times 10^{-10}$	$4.86185(3) \times 10^{-11}$	$-4.53176(3) \times 10^{-12}$	3	19
	$-1.247(1) \times 10^{-6\S}$	$2.16(1) \times 10^{-8\S}$	$-7.6(2) \times 10^{-10\S}$				
60	$-8.96719(9) \times 10^{-8}$	$5.87770(5) \times 10^{-10}$	$-9.28810(6) \times 10^{-12}$	$2.80150(1) \times 10^{-13}$	$-1.394459(6) \times 10^{-14}$	3	15
70	$-6.37956(7) \times 10^{-9}$	$1.57327(1) \times 10^{-11}$	$-1.083855(4) \times 10^{-13}$	$1.602454(5) \times 10^{-15}$	$-4.27664(2) \times 10^{-17}$	4	11
80	$-4.51561(6) \times 10^{-10}$	$4.18542(3) \times 10^{-13}$	$-1.261665(5) \times 10^{-15}$	$9.17586(2) \times 10^{-18}$	$-1.316003(4) \times 10^{-19}$	4	11
90	$-3.18908(3) \times 10^{-11}$	$1.111691(6) \times 10^{-14}$	$-1.470843(6) \times 10^{-17}$	$5.27371(1) \times 10^{-20}$	$-4.069443(7) \times 10^{-22}$	4	9
100	$-2.25081(2) \times 10^{-12}$	$2.95475(2) \times 10^{-16}$	$-1.719772(4) \times 10^{-19}$	$3.043861(3) \times 10^{-22}$	$-1.264504(1) \times 10^{-24}$	4	9

Notes: For each dimension, 16 independent simulations were run on different Intel E5 cores for two days with different random number seeds. Data were collected every 10 steps. The errors, shown in parentheses, were estimated from the standard deviations among the $q = 16$ simulations as $\sigma/\sqrt{q-1}$. The exact and best literature values (Refs. [12]*, [13]#, [20]§, [21]□) are also shown.

† $M = 0$ means that Algorithm G2 was used instead of Algorithm G3.

‡ $\times 10^{10}$: aggregated MC steps per virial coefficient spent on each regular ensemble.

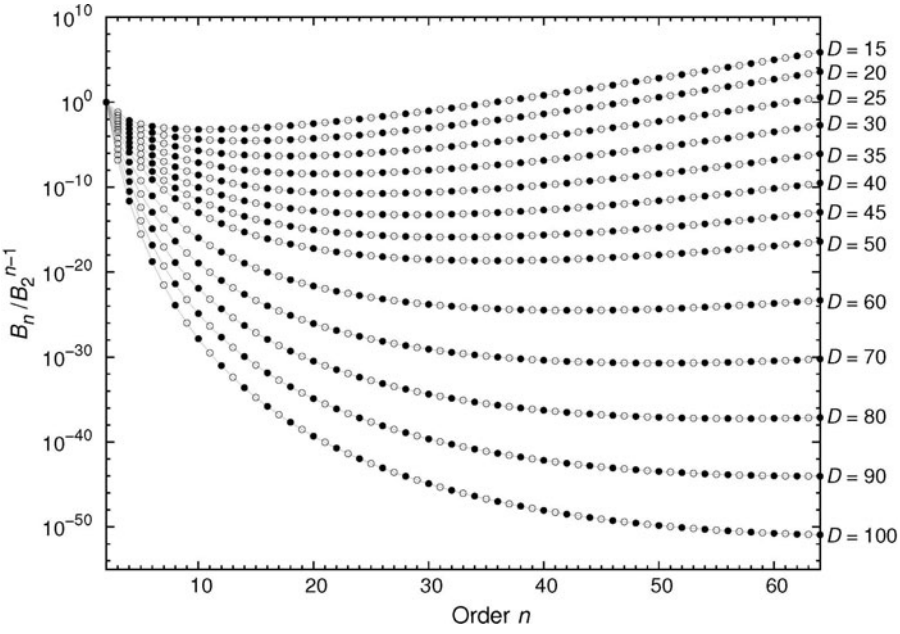

Figure 2. Virial coefficients up to the 64th order for D-dimensional ($15 \leq D \leq 100$) hard-sphere fluids. Empty and filled circles mean positive and negative figures, respectively. The thin lines are drawn to guide the eyes.

higher dimensions, for the cancellation among diagrams is less severe there.

Our method, however, was less successful in two and three dimensions. In two dimensions, only the 10th-order virial coefficient was slightly improved. The new values for the 11th to 14th virial coefficients were due to the application of Wheatley's method [18] and the dynamic hash table (Appendix 3). In three dimensions, we slightly improved the existing virial coefficients (except for $n = 12$). The relatively poor results in $n \geq 10$ were possibly due to a suboptimal implementation and shorter simulation times than those in Ref. [18]. Further, the high probability of forming biconnected configurations in the two and three dimensions diminished the main advantage of our method over the MC integration method.

To further demonstrate the advantage of our algorithm in higher dimensions, we performed a few grand ensemble simulations for $D > 8$. Since previous literature results

Table 3. High-order and high-dimensional virial coefficients.

D	$B_{16}B_2^{15}$	B_{32}/B_2^{31}	B_{48}/B_2^{47}	B_{64}/B_2^{63}	M	$T^{\dagger}$
13	$-3.3714(31) \times 10^{-3}$	$-0.7569(19)$	$-9.046(36) \times 10^{2}$	$-2.195(27) \times 10^{6}$	1	72
15	$-1.14377(38) \times 10^{-3}$	$-0.20850(8)$	$-2.6398(19) \times 10^{2}$	$-7.679(25) \times 10^{5}$	1	122
17	$-3.0401(6) \times 10^{-4}$	$-0.037331(11)$	$-42.765(23)$	$-1.2731(19) \times 10^{5}$	1	152
19	$-6.9001(8) \times 10^{-5}$	$-5.1554(10) \times 10^{-3}$	$-4.8375(14)$	$-1.3280(10) \times 10^{4}$	1	180
20	$-3.1548(8) \times 10^{-5}$	$-1.7923(7) \times 10^{-3}$	$-1.4838(8)$	$-3.814(4) \times 10^{3}$	1	46
25	$-4.8956(9) \times 10^{-7}$	$-6.1699(8) \times 10^{-6}$	$-2.3374(4) \times 10^{-3}$	$-3.6910(22)$	1	53
30	$-6.0797(7) \times 10^{-9}$	$-1.51472(17) \times 10^{-8}$	$-2.3087(5) \times 10^{-6}$	$-1.9648(6) \times 10^{-3}$	2	28
35	$-6.8782(7) \times 10^{-11}$	$-3.2415(4) \times 10^{-11}$	$-1.88739(18) \times 10^{-9}$	$-8.1952(15) \times 10^{-7}$	2	28
40	$-7.4718(9) \times 10^{-13}$	$-6.5443(6) \times 10^{-14}$	$-1.42465(18) \times 10^{-12}$	$-3.0857(8) \times 10^{-10}$	3	9.3
45	$-7.9716(5) \times 10^{-15}$	$-1.28923(8) \times 10^{-16}$	$-1.03934(11) \times 10^{-15}$	$-1.11221(21) \times 10^{-13}$	3	9.0
50	$-8.4489(5) \times 10^{-17}$	$-2.51590(16) \times 10^{-19}$	$-7.4814(5) \times 10^{-19}$	$-3.9401(4) \times 10^{-17}$	3	8.1
60	$-9.45410(33) \times 10^{-21}$	$-9.52899(29) \times 10^{-25}$	$-3.83765(16) \times 10^{-25}$	$-4.87205(39) \times 10^{-24}$	3	9.0
70	$-1.062876(27) \times 10^{-24}$	$-3.62589(4) \times 10^{-30}$	$-1.97545(5) \times 10^{-31}$	$-6.03218(21) \times 10^{-31}$	4	6.7
80	$-1.203596(12) \times 10^{-28}$	$-1.390281(16) \times 10^{-35}$	$-1.024179(7) \times 10^{-37}$	$-7.52107(14) \times 10^{-38}$	4	6.6
90	$-1.371960(7) \times 10^{-32}$	$-5.367475(34) \times 10^{-41}$	$-5.346640(32) \times 10^{-44}$	$-9.44062(12) \times 10^{-45}$	4	6.3
100	$-1.572827(8) \times 10^{-36}$	$-2.084092(8) \times 10^{-46}$	$-2.807215(8) \times 10^{-50}$	$-1.191883(6) \times 10^{-51}$	4	6.1

† $\times 10^{8}$: average MC steps per virial coefficient spent on each regular ensemble.

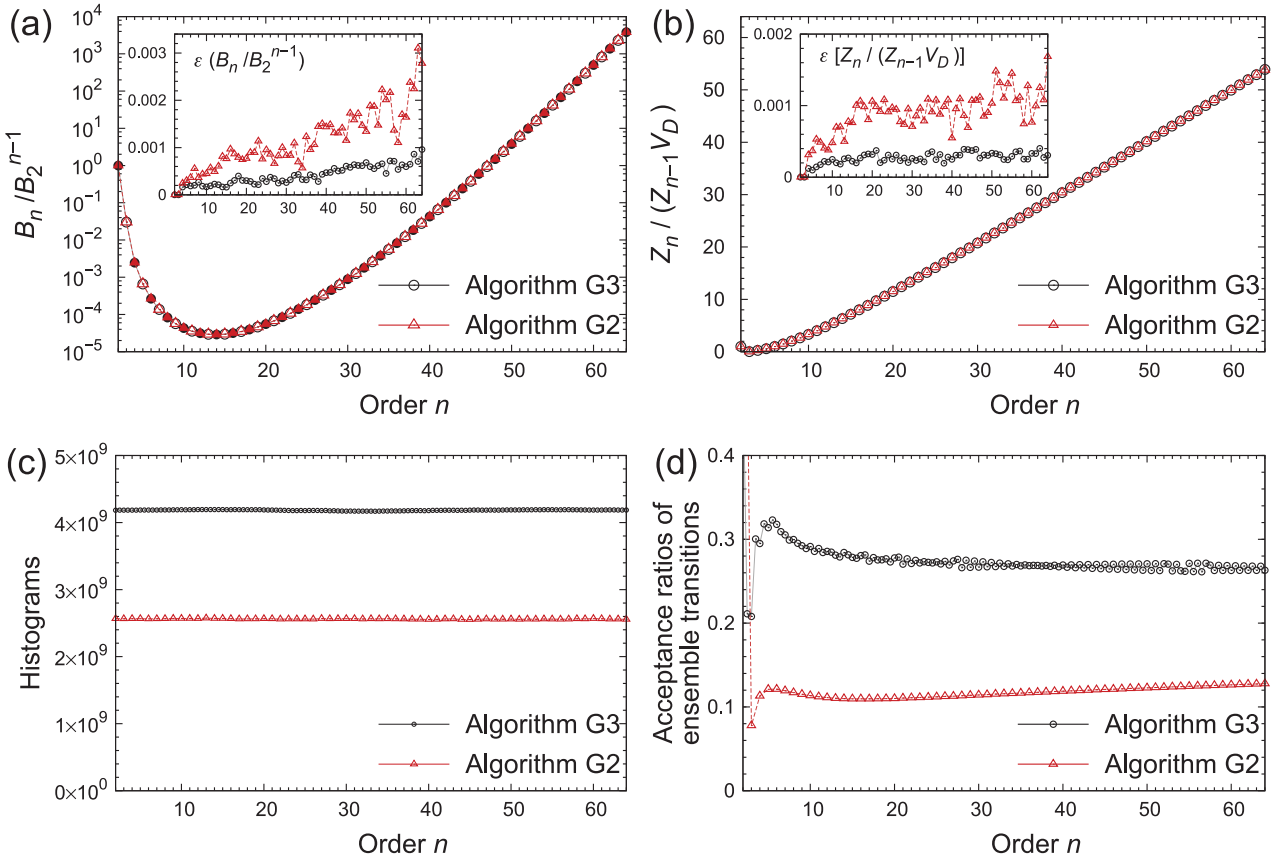

Figure 3. Comparison of Algorithms G3 ($M = 1$) and G2 for $D = 20$, $n_{\max} = 64$. (a) Virial coefficients B_n/B_2^{n-1}; inset: the relative errors. Empty and filled symbols mean positive and negative figures, respectively. (b) Inverse activities $Z_n/(Z_{n-1}V_D)$; inset: the relative errors. (c) Histograms. (d) Acceptance ratios of ensemble transitions. In (c) and (d), the values of intermediate ensembles are shown at half integer orders. The thin lines are drawn to guide the eyes. The results were obtained after a same amount of time for the two algorithms. In this case, Algorithm G3 achieved smaller errors for the virial coefficients and inverse activities, and higher transition rates among ensembles than Algorithm G2.

in these dimensions were limited to low orders, our first set of simulations were restricted to the first eight virial coefficients for comparison. We used Algorithm G2 for dimensions $D \leq 14$, otherwise Algorithm G3 (with more intermediate ensembles as the dimensionality D increases). As shown in Table 2, the precision improved with D. This behaviour was in strong contrast to the MC integration method [20], which deteriorated rapidly with increasing D and n. For example, with $D = 50$, the fourth to sixth virial coefficients obtained in a previous study using MC integration were $-1.247(1) \times 10^{-6}$, $2.16(1) \times 10^{-8}$, and $-7.6(2) \times 10^{-10}$, respectively, and higher order virial coefficients were unattainable [20]. We obtained more precise values of $-1.24719(1) \times 10^{-6}$, $2.162450(15) \times 10^{-8}$, $-7.86216(5)$

Table 4. Estimated radii of convergence.

D	n range for a_1	a_1	n range for $\rho^{(R)}$ or $\rho^{(Z)}$	$\rho^{(R)}V_D$	$\rho^{(Z)}V_D$	$\rho_{\mathrm{PY}}V_D$
7¶	[20, 64]	0.995	[8, 20]	1.458	1.485	1.3935076... [43]
9¶	[20, 64]	0.998	[8, 20]	1.227	1.260	1.2260080... [43]
11¶	[20, 64]	1.006	[20, 32]	1.148	1.143	1.1407695... [43]
13	[20, 64]	1.008	[20, 64]	1.097	1.090	1.0918875... [43]
15	[30, 64]	1.003	[30, 64]	1.065	1.065	1.0618138...
17	[30, 64]	1.003	[30, 64]	1.044	1.043	1.0424522...
19	[30, 64]	1.001	[30, 64]	1.031	1.030	1.0295924...

¶ The results in these dimensions were less reliable.

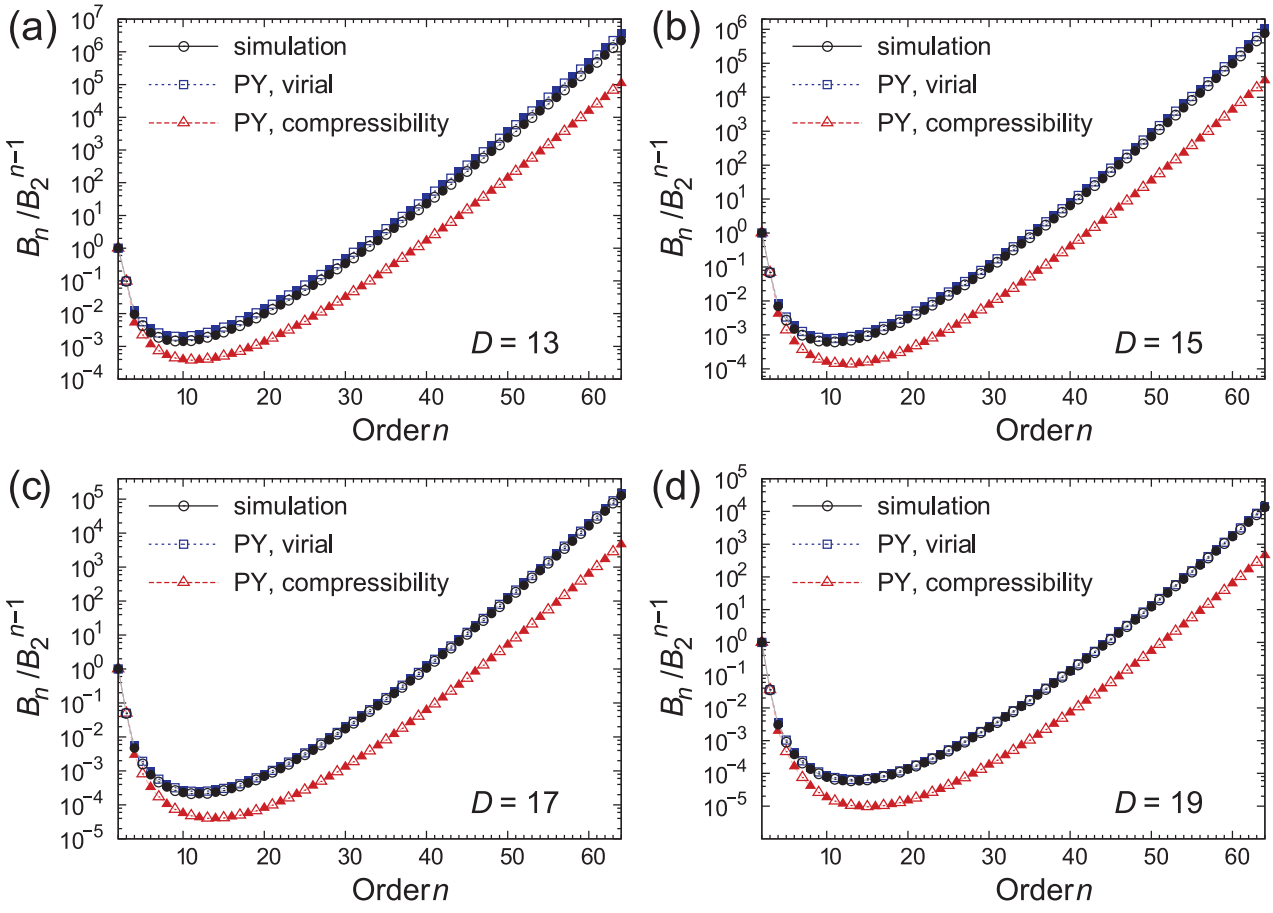

Figure 4. Virial coefficients up to the 64th order for (a) 13-, (b) 15-, (c) 17-, and (d) 19-dimensional hard-sphere fluids compared to the analytic results from the PY closure. Circles: simulation; squares: PY virial route; triangles: PY compressibility route. Empty and filled symbols mean positive and negative figures, respectively. The thin lines are drawn to guide the eyes.

$\times 10^{-10}$, respectively, and the values of the next two virial coefficients $4.86185(3) \times 10^{-11}$ and $-4.53176(3) \times 10^{-12}$ with similar relative errors.

We also computed higher order coefficients for dimensions $D > 10$. For each dimension D, we ran a grand ensemble simulation with $n_{\max} = 64$. The value 64, which is the computer word size, was chosen out of coding convenience, can be increased in the future. We exclusively used Algorithm G3, for efficiency. The virial coefficients for $D > 12$ up to 64th order are shown in Figure 2 and Table 3. As a general trend, the magnitude of the virial coefficient as a function of n first decreased to a minimum and then increased monotonically afterwards. Thus, high-order virial coefficients are essential in studying the asymptotic behaviour.

We compared the performance of Algorithms G2 and G3 in Figure 3 for the $D = 20$ case. Both algorithms obtained similar virial coefficients (Figure 3(a)) and inverse activities (Figure 3(b)), and sampled flat distributions along n (Figure 3(c)). Algorithm G3, however, achieved higher transition rates among ensembles (Figures 3(d)) and smaller errors in virial coefficients and activities (insets of Figure 3(a) and 3(b)), suggesting a higher sampling efficiency. Note that Algorithm G2 used a more elaborated pair list in the heat-bath technique, which can also contribute to the lower speed.

3.2. *Radius of convergence*

We applied Equations (22) and (23) to estimate the radii of convergence of the virial series in a few dimensions $D < 20$. Our results may have yet to reach the asymptotic regime in higher dimensions because of the small cut-off $n_{\max} = 64$. As shown in Table 4, the radii of convergence obtained from the two routes generally agreed with each other. In all cases, the coefficient $a_1 = \lim_{n\to\infty} Z_{n+1}/(nV_D Z_n)$ for the inverse activity was close to unity, and this might be exactly so. This hypothesis, however, was not enforced in estimating $\rho^{(Z)}$.

For dimensions $D > 10$, despite a considerable difference in the virial coefficients (Figure 4), the radii of convergence obtained here agreed well with those obtained from

the exact solutions [39,40–43] under the Percus–Yevick (PY) closure (Appendix 11), as shown in Table 4, supporting a previous conjecture [43]. In these dimensions, the virial coefficients from our MC calculations were closer to the results from the virial route than those from the compressibility route. The radii of convergence from the two routes, however, are expected to be the same, as they both arise from a square-root branch point on the negative axis of ρ ($D \geq 5$). It appeared that the PY solution yielded an increasingly accurate radius of convergence as the dimensionality increases. There were, however, appreciable discrepancies in seven dimensions, although the precision in this dimension was somewhat low. Even for $D > 12$, higher order virial coefficients may behave differently from the first 64 ones obtained here, and explicit virial coefficients of higher orders and precision are needed in the future.

4. Summary and discussions

In summary, we have shown that the Mayer sampling method [23,24] was useful in computing virial coefficients of high-dimensional hard-sphere fluids. The key is to find a reference integral, through either the exact ring integral [24] or the biconnected partition function obtained from the grand ensemble simulation. In this way, the average signed star content obtained from the MC trajectory can be translated to the virial coefficient. The method is superior to the conventional MC integration method in high dimensions since it naturally overcomes the difficulty of achieving the required biconnectivity, and the efficiency gain over the latter grows rapidly with the dimensionality. It is also technically simpler than the MC integration method as the enumeration of spanning trees is not required.

The new virial coefficients obtained here can be helpful for the development of fluid theories in high dimensions [22,42–44]. We, however, have yet to reach the limit of the method, especially in high dimensions. Extensions to higher order virial coefficients may benefit from more advanced simulation techniques [34,45] and parallel implementations based on graphics processors [26].

Acknowledgements

It is a pleasure to thank Dr Chun-Liang Lai for helpful discussions. The authors are also grateful for the support of the Texas Advanced Computing Center at the University of Texas at Austin and that of the National Institute for Computational Sciences at the University of Tennessee for the computational time on the Stampede [46] and Kraken [47] supercomputers, respectively (TG-MCA93S001).

Funding

The authors gratefully acknowledge the financial support of the National Science Foundation [grant number CHE-1152876]; the Robert A. Welch Foundation [grant number H-0037].

Supplemental data

Supplemental data for this article can be accessed here. Additional data and computer code are available upon request [50].

References

[1] J.-P. Hansen and I.R. McDonald, *Theory of Simple Liquids*, 3rd ed. (Academic Press, Amsterdam, 2007).
[2] D.A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
[3] L.E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980).
[4] R. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[5] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[6] B.J. Alder and T.E. Wainwright, J. Chem. Phys. **27**, 1208 (1957).
[7] J.D. van der Waals, Proc. Kon. Akad. Wetensch. Amster. **1**, 138 (1899); J.J. van Laar, Proc. Kon. Akad. Wetensch. Amster. **1**, 273 (1899); L. Boltzmann, Proc. Kon. Akad. Wetensch. Amster. **1**, 398 (1899); J.S. Rowlinson, Mol. Phys. **7**, 593 (1964); P.C. Hemmer, J. Chem. Phys. **42**, 1116 (1965); C.G. Joslin, J. Chem. Phys. **77**, 2701 (1982).
[8] B.R.A. Nijboer and L. Van Hove, Phys. Rev. **85**, 777 (1952); G.S. Rushbrooke and P. Hutchinson, Physica **27**, 647 (1961); S. Kim and D. Henderson, Phys. Lett. A **27**, 378 (1968).
[9] J.E. Kilpatrick, Adv. Chem. Phys. **20**, 39 (1971).
[10] K. Kratky, J. Stat. Phys. **27**, 533 (1982).
[11] M. Luban and A. Baram, J. Chem. Phys. **76**, 3233 (1982).
[12] N. Clisby and B.M. McCoy, J. Stat. Phys. **114**, 1343 (2004).
[13] I. Lyberg, J. Stat. Phys. **119**, 747 (2005).
[14] F.H. Ree and W.G. Hoover, J. Chem. Phys. **40**, 939 (1964).
[15] F.H. Ree and W.G. Hoover, J. Chem. Phys. **41**, 1635 (1964).
[16] N. Clisby and B.M. McCoy, J. Stat. Phys. **122**, 15 (2006).
[17] S. Labík, J. Kolafa, and A. Malijevský, Phys. Rev. E **71**, 021105 (2005).
[18] R.J. Wheatley, Phys. Rev. Lett. **110**, 200601 (2013).
[19] M.N. Rosenbluth and A.W. Rosenbluth, J. Chem. Phys. **22**, 881 (1954); F.H. Ree and W.G. Hoover, J. Chem. Phys. **40**, 2048 (1964); **46**, 4181 (1967); M. Bishop, A. Masters, and J.H.R. Clarke, J. Chem. Phys. **110**, 11449 (1999); A.Y. Vlasov, X.M. You, and A.J. Masters, Mol. Phys. **100**, 3313 (2002); E.J. Janse van Rensburg, J. Phys. A **26**, 4805 (1993); M. Bishop, A. Masters, and A.Y. Vlasov, J. Chem. Phys. **121**, 6884 (2004); **122**, 154502 (2005).
[20] N. Clisby and B.M. McCoy, J. Stat. Phys. **114**, 1361 (2004).
[21] M. Bishop, N. Clisby, and P.A. Whitlock, J. Chem. Phys. **128**, 034506 (2008).
[22] J.G. Loeser, Z. Zhen, S. Kais, and D.R. Herschbach, J. Chem. Phys. **95**, 4525 (1991).
[23] S. Rast, P.H. Fries, and H. Krienke, Mol. Phys. **96**, 1543 (1999); S. Labík, H. Gabrielová, J. Kolafa, and A. Malijevský, *ibid.* **101**, 1139 (2003).
[24] J.K. Singh and D.A. Kofke, Phys. Rev. Lett. **92**, 220601 (2004).
[25] S.K. Kwak and D.A. Kofke, J. Chem. Phys. **122** (2005); K.M. Benjamin, A.J. Schultz, and D.A. Kofke, Ind. Eng. Chem. Res. **45**, 5566 (2006); K.M. Benjamin, J.K. Singh, A.J. Schultz, and D.A. Kofke, J. Phys. Chem. B **111**, 11463 (2007); K.M. Benjamin, A.J. Schultz, and D.A. Kofke, J. Phys. Chem. C **111**, 16021 (2007); J. Phys. Chem. B **113**, 7810 (2009); A.J. Schultz and D.A. Kofke, Mol. Phys. **107**, 2309 (2009); J. Chem. Phys. **130**, 224104 (2009); D.J. Naresh and J.K. Singh, Fluid Phase Equilib. **279**, 47 (2009);

285, 36 (2009); T.B. Tan, A.J. Schultz, and D.A. Kofke, Mol. Phys. **109**, 123 (2010); K.R.S. Shaul, A.J. Schultz, and D.A. Kofke, Mol. Simul. **36**, 1282 (2010); Mol. Phys. **109**, 2395 (2011); J. Chem. Phys. **137** (2012); J.H. Yang, A.J. Schultz, J.R. Errington, and D.A. Kofke, J. Chem. Phys. **138** (2013); H.M. Kim, A.J. Schultz, and D.A. Kofke, Fluid Phase Equilib. **351**, 69 (2013).
[26] A.J. Schultz, N.S. Barlow, V. Chaudhary, and D.A. Kofke, Mol. Phys. **111**, 535 (2013).
[27] B.D. McKay, Congr. Numer. **30**, 45 (1981).
[28] R.E. Tarjan, Discrete Math. **55**, 221 (1985).
[29] D. Rose, R. Tarjan, and G. Lueker, SIAM J. Comput. **5**, 266 (1976).
[30] R.E. Tarjan and M. Yannakakis, SIAM J. Comput. **13**, 566 (1984).
[31] A. Berry, J.R.S. Blair, and P. Heggernes, *Graph-Theoretic Concepts in Computer Science* (Springer, Berlin Heidelberg, 2002).
[32] Mathematica® (Wolfram Research, Champaign, IL, 2012).
[33] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, San Diego, CA, 2002).
[34] M.E.J. Newman and G.T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, Oxford, 1999).
[35] A.F. Voter, J. Chem. Phys. **82**, 1890 (1985); M.A. Miller and W.P. Reinhardt, J. Chem. Phys. **113**, 7035 (2000); C. Jarzynski, Phys. Rev. E **65**, 046122 (2002); S. Vaikuntanathan and C. Jarzynski, Phys. Rev. Lett. **100**, 190601 (2008); J. Chem. Phys. **134**, 054107 (2011).
[36] C.H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[37] N.S. Barlow, A.J. Schultz, S.J. Weinstein, and D.A. Kofke, J. Chem. Phys. **137**, 204102 (2012).
[38] M. Ončák, A. Malijevský, J. Kolafa, and S. Labík, Condens. Matter Phys. **15**, 23004 (2012).
[39] M.S. Wertheim, Phys. Rev. Lett. **10**, 321 (1963); J. Math. Phys. **5**, 643 (1964); E. Thiele, J. Chem. Phys. **39**, 474 (1963).
[40] R.J. Baxter, Aust. J. Phys. **21**, 563 (1968).
[41] E. Leutheusser, Physica A **127**, 667 (1984).
[42] M. Robles, M.L. de Haro, and A. Santos, J. Chem. Phys. **120**, 9113 (2004).
[43] R.D. Rohrmann, M. Robles, M.L. de Haro, and A. Santos, J. Chem. Phys. **129**, 014510 (2008).
[44] H.L. Frisch, N. Rivier, and D. Wyler, Phys. Rev. Lett. **55**, 550 (1985); D. Wyler, N. Rivier, and H.L. Frisch, Phys. Rev. A **36**, 2422 (1987); H.L. Frisch and J.K. Percus, Phys. Rev. E **60**, 2942 (1999); M. Robles, M.L. de Haro, and A. Santos, J. Chem. Phys. **126**, 016101 (2007); R.D. Rohrmann and A. Santos, Phys. Rev. E **76**, 051202 (2007); M. Adda-Bedia, E. Katzav, and D. Vella, J. Chem. Phys. **128**, 184508 (2008); **129**, 144506 (2008); G. Parisi and F. Slanina, Phys. Rev. E **62**, 6554 (2000).
[45] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, San Diego, CA, 2002).
[46] Stampede. Available from http://www.tacc.utexas.edu/stampede/.
[47] Kraken. Available from http://www.nics.tennessee.edu/computing-resources/kraken.
[48] T.H. Cormen, *Introduction to Algorithms*, 3rd ed. (MIT Press, Cambridge, MA, 2009).
[49] B.D. McKay and A. Piperno, J. Symbolic Comput. **60**, 94 (2014); S.G. Hartke and A.J. Radcliffe, *Communicating Mathematics* (American Mathematical Society, Providence, RI, 2009).
[50] http://bmb.utmb.edu/pettitt/software/. 6.
[51] T. Ohtsuki, SIAM J. Comput. **5**, 133 (1976); D. Dahlhaus, *Graph-Theoretic Concepts in Computer Science* (Springer, Berlin, Heidelberg, 1997), vol. 1335, Chap. 13, pp. 132–143; B.W. Peyton, SIAM J. Matrix Anal. Appl. **23**, 271 (2001).
[52] R. Bellman, J. ACM **9**, 61 (1962); M. Held and R. Karp, J. Soc. Ind. Appl. Math. **10**, 196 (1962).
[53] S. Kohn, A. Gottlieb, and M. Kohn, presented at the Proceedings of the 1977 Annual Conference, ACM, New York, 1977 (unpublished); E.T. Bax, Inf. Process. Lett. **47**, 203 (1993); A. Björklund, T. Husfeldt, P. Kaski, and M. Koivisto, *Automata, Languages and Programming* (Springer, Berlin, Heidelberg, 2008).
[54] R.M. Karp, Oper. Res. Lett. **1**, 49 (1982).
[55] A. Björklund, presented at the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society, New York, 2010 (unpublished).
[56] G.H. Hardy, E.M. Wright, D.R. Heath-Brown, and J.H. Silverman, *An Introduction to the Theory of Numbers*, 6th ed. (Oxford University Press, Oxford, 2008); P. Flajolet and R. Sedgewick, *Analytic Combinatorics* (Cambridge University Press, Cambridge, 2009).
[57] D.S. Dummit and R.M. Foote, *Abstract Algebra*, 3rd ed. (Wiley, Hoboken, NJ, 2004).
[58] W. Bosma, J. Cannon, and C. Playoust, J. Symb. Comput. **24**, 235 (1997).

Appendix 1. Wheatley's method for the signed star content

Here we review Wheatley's method [18] of computing the signed star content. Below, we assume that the input RH diagram G has n vertices, and denote the set of vertices as $V = \{1, 2, \ldots, n\}$.

First, recall that the Boltzmann weight

$$f_Q(V) = \prod_{i,j \in V} (1 + f_{ij}) = \sum_H \prod_{\{i,j\} \in E(H)} f_{ij} \tag{A1}$$

is the sum of the integrands $\prod_{\{i,j\} \in E(H)} f_{ij}$ (hereinafter called *f-products*) of Mayer diagrams H of n vertices, with an arbitrary edge set $E(H)$. As shown in Section 2.2, the signed star content is equal to the partial sum $f_B(V)$ of the *f*-products of only *biconnected* Mayer diagrams for a hard-sphere fluid. More generally, we define $f_Q(S)$ and $f_B(S)$ for an arbitrary subset S ($S \subseteq V$). The method computes $f_B(S)$ from the 2^n values of $f_Q(S)$.

We first compute the sum f_C of connected diagrams for a subset S ($S \subseteq V$) by removing the *f*-products of disconnected diagrams using the following relation [18]:

$$f_C(S) = f_Q(S) - \sum_{\substack{T \subset S, \\ v_0 \in T}} f_C(T) f_Q(S \backslash T), \tag{A2}$$

where v_0 is the lowest vertex in S, and the sum is carried over proper subsets T of S containing v_0. The notation $S \backslash T$ means the set of vertices that belong to S but not to T. In practice, we compute and save the $f_C(S)$ values in ascending order of $|S|$ to ensure that $f_C(T)$ on the right side of (A2) are available when evaluating $f_C(S)$, as $|T| < |S|$.

The second recursion further removes the *f*-products of Mayer diagrams of articulation points. We define $f_{B,v}(S)$ as the sum over connected subgraphs of vertices in the subset S without an

articulation point lower than v [18], and

$$f_{B,v+1}(S) = f_{B,v}(S) - \sum_{\substack{T_v \subset S,\\ v,v_1 \in T_v}} f_{B,v+1}(T_v) f_{B,v}(T_v^\dagger), \qquad \text{(A3)}$$

where v_1 is the lowest vertex in $S\backslash\{v\}$, and $T_v^\dagger = (S\backslash T) \cup \{v\}$. For efficiency, we compute and save the $f_{B,v}(S)$ values in ascending order of v, or $|S|$ for the same-v cases. In this way, quantities on the right side of (A3) are available when we evaluate $f_{B,v+1}(S)$ (because $|T_v|, |T_v^\dagger| < |S|$). In the $v = 1$ case, we have $f_{B,1} = f_C$. The final result is $f_B(V) = f_{B,n+1}(V)$. Since Equation (A3) involves $f_{B,v}$ of only two successive v and $v + 1$, we can discard all $f_{B,v'}$ values of $v' < v$ to reduce the storage from $n \cdot 2^n$ [18] to $2 \cdot 2^n$ values.

Finally, we can skip some subsets S in the above recursions. Only (Mayer) subgraphs H of the input RH diagram G contribute non-zero f-products to the sum in Equation (A1), hence to $f_C(V)$ and $f_{B,v}(V)$. Similarly, $f_C(S)$ and $f_{B,v}(S)$ only include subgraphs of the induced subgraph $G(S)$. Thus, if $G(S)$ is disconnected, then by definition $f_C(S) = 0$, and $f_{B,v}(S) = 0$ for any v. Similarly, if $G(S)$ is connected and the lowest articulation point is v, then $f_{B,u}(S) = 0$ for $u > v$. The tests are cheap for the complexity is polynomial while that of Equations (A2) and (A3) is of order $2^{|S|}$.

Appendix 2. Reduced lookup table for the star and ring contents

Here we review the technique in Ref. [17], which extends the usual lookup table technique for computing the star content to the ninth order. In the modification below, the lookup table saves both the star and the ring contents.

Given an n-vertex RH diagram, we first find a chain of $k < n$ consecutive f-bonds v_1–v_2–$\cdots$–v_{k+1} by a depth-first search. If no such chain exists, we compute the star content directly and set the ring content to zero (for a non-zero ring content requires at least one ring, hence chain, subgraph). Otherwise, we reorder the n vertices such that $v_1, v_2, \ldots, v_{k+1}$ are the first $k + 1$ ones, and arrange the rest of the vertices arbitrarily. The collective state of the $n(n-1)/2 - k$ undetermined edges (whether each edge is an e-bond or an f-bond) then serves as the key to the reduced lookup table.

Unlike the original version [17], we fill the lookup table on the fly. Once the star and ring contents of an unknown diagram are computed, we enumerate all vertex orderings that allow the first $k + 1$ vertices to form a chain. The same star and ring contents are then assigned to all resulting diagrams, which reshuffle the undetermined edges.

The value $k = 4$ used by Labík *et al.* (which ensures a complete coverage of all biconnected diagrams) [17] can be relaxed. Even with $k = 8$, the lookup table still covers over 99.99% diagrams in a trajectory for $n = 9$ and $D = 2$. We used $k = 7$ for $n = 9$. The method can be stretched to compute the 10th-order virial coefficients on computers with large memory [46]. We used $k = 8$ in some simulations for $n = 10$ and $D = 2$–5. While this technique is easier to program, the alternative in Appendix 3 is more general.

Appendix 3. Dynamic hash table for the star and ring contents

Here we extend the hash table technique by Clisby and McCoy [16] to compute the star and ring contents of diagrams of orders $n > 10$. The basic use of the hash table [48] for this purpose is the following. Given an n-vertex RH diagram, we first save the collective state of all edges in an integer c of $n(n-1)/2$ bits. We then map c to another integer s with fewer m bits by a hash function [48], and use s as the *index* of the hash table. Each entry of the hash table is a collection of the longer integers c that share the same index s, along with the star contents (and the ring contents, if needed). The values m typically varied from 24 to 28 according to the memory limit.

We can reduce the size of the hash table by first mapping the input diagram to a representative one. As in Ref. [16], we usually use the program NAUTY [27,49] to map a group of isomorphic graphs to a single graph called the *canonical label*. Since the program was somewhat slow [16], we used a less representative but easier-to-find diagram for $n = 9$ and 10. This diagram, hereinafter called a *degree representative*, sorts vertices by the degrees, but leaves the orderings of vertices with the same degrees undetermined (see Figure 5 for an example).

The previous implementation employed a static hash table and pre-computed the star contents of all non-isomorphic biconnected diagrams [16]. For higher orders, it is more memory-efficient to fill the hash table on the fly. That is, once we find a new diagram in the trajectory, we compute the star and ring contents and save the values in the hash table. The dynamic hash table only covers most frequently occurring diagrams, and omits rare but geometrically-permissible ones (e.g., the fully connected diagram in high dimensions). Thus, it needs much less memory, and is applicable to orders $n > 10$.

In Algorithm Z (which requires no ring content), we can further save memory by limiting the hash table to diagrams free of clique separators [16], because many diagrams in the trajectory have clique separators and hence zero star contents. This, however,

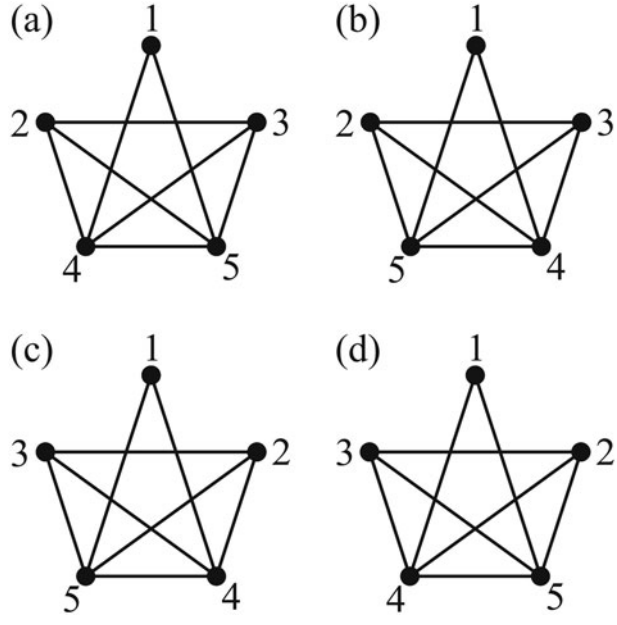

Figure 5. Four vertex orderings compatible with the degree sequence. Any of the four diagrams can be the canonical label. The degrees of the five vertices are 2, 3, 3, 4, and 4. To be compatible with the degree sequence, the degree-2 vertex must precede other vertices, and the two degree-3 vertices must precede the two degree-4 vertices. The two degree-3 vertices, however, can exchange their orders, so can the two degree-4 vertices.

requires an efficient algorithm to test clique separators (Appendix 5). Finally, if we use the degree representative, on encountering a new diagram G, we apply the star and ring contents of G to similar degree representatives G_2, which differ from G by the orders of vertices with the same degrees. For example, if G is the diagram in Figure 5(a), G_2 can be any diagram in Figure 5(b)–(d).

Appendix 4. Zero star contents of diagrams with clique separators

The star content of a diagram with clique separator(s) is zero. While the result was quoted several times [9,16,18] and some special cases were shown in Refs. [14,15], the general proof by Kilpatrick, Ford and Yu was unfortunately not published [9]. Thus, the following proof may be helpful.

We first give a few definitions. Given a diagram G of vertex set V and edge set E, and a subset of vertices U, $G(U)$ denotes the subgraph of G induced by U whose edge set is $E(U) \equiv \{\{u, v\} \in E | u, v \in U\}$. For a subgraph H, $H \cup E'$ means the subgraph obtained from adding edges in E' into H, $H \backslash E'$ means that obtained from removing the edges in E' from H, and $H \backslash V'$ means that obtained from removing the vertices in V' and the incident edges from H.

Given a clique separator C of G, and a biconnected subgraph H of G, we define the *filled graph* as $W_C[H] \equiv H \cup E(C)$, and the *reduced graph* as $M_C[H] = H \backslash E(C)$. Then:

(1) if $M_C[H] = M_C[H']$, then $W_C[H] = W_C[H']$, and vice versa;
(2) $M_C[H]$ is a subgraph of H;
(3) H is a subgraph of $W_C[H]$;
(4) $W_C[H]$ is a subgraph of G;
(5) $W_C[H]$ is biconnected;
(6) C is a clique separator of $W_C[H]$.

Properties (1)–(4) follow from the definitions. Property (5) follows from the biconnectivity of H and property (3). Property (6) follows from property (4) and that C is a clique separator of G.

We say that a clique separator C of G is *minimal* if no proper subset of C disconnects G. Then,

Lemma 1: *If C is a minimal clique separator of $W \equiv W_C[H]$, and let Y be an arbitrary union of the vertex sets $X_1, X_2, \ldots$ of the connected components of $W \backslash C$, then the induced subgraph $M(C' \cup Y)$ is connected, where $M \equiv M_C[H]$, and C' is a non-empty subset of C.*

Proof: First, in W, any vertex v in C is adjacent to at least one vertex in every X_i ($i = 1, 2, \ldots$), for, otherwise, $C \backslash \{v\}$ also separates some X_i from the rest of $W \backslash C$, and thus is a clique separator smaller than C, contradicting the minimality of C [28]. As the edges not in $E(C)$ are common in W and M, the induced graphs $M(\{v\} \cup X_i)$, and hence $M(\{v\} \cup Y)$, are connected. Obviously, the argument holds for other vertices $v', v'', \ldots$, in C, and $M(C' \cup Y)$ is connected for an arbitrary $C' = \{v, v', v'', \ldots\}$. □

Lemma 2: *If H is a biconnected subgraph of G, and C ($|C| \geq 2$) is a minimal clique separator of $W \equiv W_C[H]$, then $M \equiv M_C[H]$ is biconnected.*

Proof: We will show that M is connected and free of articulation points. First, M is connected by using $C' = C$ and $Y = \bigcup_i X_i$ in Lemma 1. Similarly, the case of $C' = C \backslash \{v\}$ and $Y = \bigcup_i X_i$ shows that any vertex v ($v \in C$) is not an articulation point. Finally, for a vertex x ($x \in X_i$), the induced graph $M(Q)$ with $Q = C \cup \bigcup_{j \neq i} X_j$ is connected. Suppose x is an articulation point of M, then Q, hence C, are entirely contained in one connected component of $M \backslash \{x\}$. Since W differs from M only by the edges in $E(C)$, x is also an articulation point of W. However, by property (5), W is biconnected and has no articulation point, a contradiction. □

Lemma 3: *For a diagram G with a clique separator C, all biconnected subgraphs H of G sharing the same filled graph $W = W_C[H]$ can be paired such that any pair contains two subgraphs that differ by a single edge.*

Proof: We prove this by an induction on the size $k = |C|$ of the clique separator C. First, with $k = 1$, C is an articulation point, and G has no biconnected subgraph. The statement holds trivially.

Now, assuming the statement holds for k up to K ($K \geq 1$), we will prove the $k = K + 1$ case. By property (3), we may consider only the biconnected subgraphs H' of W, instead of the biconnected subgraphs H of G, and show that among the former, the subgraphs H' with $W_C(H') = W$ are properly paired. By property (6), C is a clique separator of W. Consider two cases: (1) C is a minimal clique separator of W or (2) C is not. In case (1), $M \equiv M_C[H]$ is biconnected by Lemma 2. All subgraphs H' with $W_C[H'] = W$, and hence with $M_C[H'] = M$ by property (1), are biconnected by property (2). Thus, for any edge $e \in E(C)$, a subgraph H' of edge e can be paired to $H'' = H' \backslash \{e\}$ (which satisfies $W_C[H''] = W$) and the pairing covers all subgraphs of W, hence those of G, with the same filled graph W exactly once.

In case (2), W has a smaller clique separator C' ($C' \subset C$) with $|C'| \leq K$. By the induction hypothesis, all biconnected subgraphs H' of W with the same $W_{C'}[H']$ can be properly paired. Since the subgraphs H' with the same $W_{C'}[H']$ differ only by the edges in $E(C')$, their filled graphs $W_C[H']$ are the same. Thus, all biconnected subgraphs H' of W with $W_C[H'] = W$ are properly paired. □

Theorem 4: *If a diagram G has a clique separator C, the star content is zero.*

Proof: We first partition biconnected subgraphs H of G into mutually exclusive sets with each set collecting subgraphs with the same $W_C[H]$. By Lemma 3, all biconnected subgraphs in each set can be paired such that the two subgraphs in a pair differ by a single edge. However, the two subgraphs in a pair contribute $+1$ and -1 to the star content by definition, so the total contribution is zero. □

Appendix 5. Detection of clique separators

Here we briefly discuss Tarjan's algorithm for detecting clique separators in a given diagram G [28]. The key is the use of a minimal (elimination) ordering of the vertices, as explained below. If, hypothetically, we perform a Gaussian elimination on the 0–1 adjacency matrix according to some vertex ordering, the set of non-zero elements created in the process (assuming no accidental cancelling) is called the *fill-in*. The ordering is *minimal* if no proper subset of the fill-in can be produced by another ordering [29]. We can then construct an augmented diagram $F(G)$ with additional edges corresponding to the fill-in elements. Now, if we eliminate vertices in $F(G)$ according to the minimal ordering, there is always a vertex v adjacent to a clique separator of G (if G has one) right before v is eliminated [28]. We can then detect clique separators by a linear scan.

A minimal ordering can be computed using the Rose–Tarjan–Lueker algorithm (LEX-M) based on lexicographic search [29] or using the simplification MCS-M by Berry, Blair and Heggernes [31] based on maximal cardinality search (MCS) [30] among other choices [51]. Note that, unlike MCS-M, the original version MCS-P [30] does not necessarily produce a minimal ordering

and thus leaves some diagrams with clique separators undetected [16], although MCS-P runs faster than LEX-M. While MCS-M is more complex than MCS-P, it yields the fill-in (which is needed by Tarjan's algorithm) along with the minimal ordering, whereas MCS-P requires a separate step to do so. Thus, MCS-M appeared to be the best among the three for the purpose.

Appendix 6. Computation of the ring content

Here we give some algorithms to compute the ring content for dense diagrams. First, we define a few terms. Given an RH diagram G of vertex set V ($n = |V|$) and edge set E, a *closed walk* is a circular path of f-bonds in G. A (*Hamiltonian*) *cycle* is a closed walk that visits every vertex exactly once. So, the ring content $R(G)$ is equal to the number of cycles in G. There are many ways to count cycles [52,53,54,55], and the algorithms below are among the simplest.

We first review the classic algorithm by dynamic programming [52]. Let $Q_s(t, S)$ ($s, t \notin S$) be the number of ways to go from s to t visiting every vertex in S exactly once. Then, the ring content $R = Q_s(s, V\backslash\{s\})$. We compute $Q_s(t, S)$ from the following recursion:

$$Q_s(t, S) = \sum_{\tau \in S, \{\tau, t\} \in E} Q_s(\tau, S\backslash\{\tau\}), \tag{F1}$$

which expresses the extension of a path from s to τ via vertices in $S\backslash\{\tau\}$ by an f-bond $\{\tau, t\}$. For efficiency, we compute $Q_s(t, S)$ in ascending order of $|S|$ to ensure every term on the right side of Equation (F1) is available (much like in Wheatley's method in Appendix 1). In practice, we prefer this method to that discussed in Section 2.6 for a diagram of no less than $2.5n$ f-bonds because the time complexity of this method is $N_f 2^{n-1}$ instead of $n!$ (where N_f is the number of f-bonds in G).

Next, we give an algorithm to compute the ring content $R(G)$ for a nearly complete diagram. We denote by $\tilde{E}$ the set of e-bonds in G, and allow e-bonds in closed walks and cycles below. Since $R(G)$ counts cycles of no e-bond, we can start with all cycles, and gradually remove those that pass through some subsets of the e-bonds.

Consider a diagram G of two adjacent e-bonds: $\tilde{E} = \{e_1, e_2\}$. We start with the $(n-1)!/2$ unrestricted cycles, then subtract the $(n-2)!$ cycles that pass through e_1, and do the same for e_2, and, finally, add back the $(n-3)!$ cycles that pass through both e_1 and e_2 to correct the double counting. So, $R(G) = (n-1)!/2 - (n-2)! \times 2 + (n-3)!$.

For a general G, we apply the inclusion–exclusion principle [56], and

$$R(G) = \sum_{B \subseteq \tilde{E}} (-1)^{|B|} N(n, B), \tag{F2}$$

where $N(n, B)$ is the number of cycles that pass through all e-bonds in a subset B of $\tilde{E}$, and $|B|$ is the number of e-bonds in B. If three or more e-bonds in B meet at a single vertex, then they cannot be traversed by a cycle and $N(n, B) = 0$. A cycle is also impossible if the edges in B form a closed walk of fewer than n vertices. Thus, e-bonds in B can only form either a set of disjoint paths (or *segments* below) or a cycle, and

$$N(n, B) = \begin{cases} 1 & \text{if the edges in } B \text{ form a cycle,} \\ (n - |B| - 1)! 2^{M(B)-1} & \text{otherwise,} \end{cases} \tag{F3}$$

where $M(B)$ is the number of segments in B. Here $(n - |B| - 1)!$ gives the number of circular permutations of the free vertices in B (each e-bond fixes a vertex, making it unfree), and there are $2^{M(B)-1}$ ways of changing the directions of all segments but the first, which fixes the direction of the cycle. For example, for $B = \{e_1, e_2\}$, if e_1 and e_2 are adjacent, then $M(B) = 1$ and $N(n, B) = (n-3)!$; otherwise, $M(B) = 2$, and $N(n, B) = (n-3)!\,2$.

Algorithmically, we can enumerate all qualified sets B of e-bonds by a depth-first search (similar to the direct method for the star content in Section 2.4). In practice, we used this algorithm only for diagrams that differed from the complete diagram by no greater than $1.5n - 3$ edges.

We now define the ratio of $R(G)$ to the leading term $N(n, \emptyset) = (n-1)!/2$ of Equation (F2):

$$r(n, \tilde{E}) \equiv \frac{R(G)}{(n-1)!/2} = \sum_{B \subseteq \tilde{E}} \frac{(-1)^{|B|} 2^{M(B)}}{(n-1)^{(-)}_{|B|}},$$

where $(x)^{(-)}_m = x(x-1)\cdots(x-m+1)$ is the falling Pochhammer symbol. This definition allows a 'convolution' on $r_n(\tilde{E})$ if $\tilde{E}$ is the union of two disjoint edge sets $\tilde{E}_1$ and $\tilde{E}_2$:

$$r(n, \tilde{E}_1 \cup \tilde{E}_2) = \sum_{B_1 \subseteq \tilde{E}_1} \frac{(-1)^{|B_1|} 2^{M(B_1)}}{(n-1)^{(-)}_{|B_1|}} r(n - |B_1|, \tilde{E}_2). \tag{F4}$$

Equations (F2)–(F4) yield a few compact expressions in special cases. The ring content of k disjoint e-bonds is ${}_1F_1(-k; -n+1; -2)\,(n-1)!/2$, where ${}_1F_1(a; b; z) = \sum_{l=0}^{\infty} \frac{(-a)^{(-)}_l}{(-b)^{(-)}_l} \frac{z^l}{l!}$ is the confluent hypergeometric function. For a two-e-bond segment with k disjoint e-bonds, the ring content is ${}_1F_1(-k-2; -n+1; -2)\,(n-1)!/2 - {}_1F_1(-k; -n+3; -2)\,(n-3)!$, etc. The ring content of a k-vertex clique of e-bonds is $(n-k-1)^{(-)}_{k-1}(n-k)!/2$, which vanishes if $n < 2k$. In the last case, there are $(n-k-1)^{(-)}_{k-1}$ ways of arranging the k vertices of the clique on the cycle without any two of them being neighbours, $(n-k)!$ arrangements for the rest of the vertices, and the $1/2$ fixes the cycle direction.

Finally, we mention an algorithm that uses the inclusion–exclusion principle differently [53,54]. Let $A_s(n, S)$ be the number of length-n closed walks of f-bonds through a fixed vertex s and vertices in S (a vertex may be visited several times in the walk). Then, the ring content is

$$R = \sum_{S \subseteq V, s \in S} (-1)^{|V\backslash S|} A_s(n, S),$$

where a closed walk contributes $+1$ to the sum if it is a cycle or 0 otherwise, because of the alternating sign $(-1)^{|V\backslash S|}$. Then, we can compute $A_s(n, S)$ by counting the number of paths of length k, $A_{s,S}(k, t)$, that start from vertex s, and end at vertex t, using a recursion over k for fixed s and S, and $A_s(n, S) = A_{s,S}(n, s)$ [54]. This algorithm does not perform as well as the first algorithm in this appendix. It, however, uses only $O(n)$ memory [54], and thus may be useful when the memory is limited.

Appendix 7. Ergodicity of sampling biconnected configurations

Here, we show that the sampling of biconnected configurations using the single-point move defined in Section 2.7 is ergodic. We do so by constructing a path of a finite number of steps that starts

from a biconnected configuration to the fully collapsed configuration $\mathbf{r}_1 = \cdots = \mathbf{r}_n$. Thus, any two biconnected configurations are connected by a path of a finite number of MC steps via the fully collapsed configuration [34].

First, if the configuration is fully connected, the distance $d(i, j) \leq 1$ for any pair $\{i, j\}$. Then, we may move all points to $\mathbf{r}_1$ without breaking the full connectivity and our statement holds trivially. Below, we will show that a not fully connected configuration can be transformed to a fully connected one in a finite number of steps.

Let $\{i, j\}$ be the pair of points with the longest distance $c = d(i, j)$, and $c > 1$. We first show the following lemma: we can move point j towards point i for a distance no less than $1/(2c)$ without breaking the biconnectivity of the configuration. Our move will keep the configuration biconnected by preserving all existing f-bonds $\{j, k\}$. Consider an arbitrary point k $(k \neq i, j)$. Denote $a \equiv d(i, k)$ and $b \equiv d(k, j)$ for the values before the move. As c is the longest pair distance, we have $c \geq a$, $c \geq b$, and the angles $\angle jik$ and $\angle ijk$ are acute. Consider the three cases in Figure 6. (A) If $a \leq b$, we can move j to any place on the line segment i–j without increasing $d(k, j)$. Similarly, if both a and b are less than 1, $d(k, j)$ cannot exceed 1.0 under the same move. Thus, the maximal permissible distance of the move is $c > 1/(2c)$ in both cases. (B) If $a > b \geq 1$, $d(k, j)$ decreases monotonically as j moves towards i up to the projection h of k on the line segment i–j. The maximal permissible distance $d(j, h)$ is $(b^2 + c^2 - a^2)/(2c) \geq b^2/(2c) \geq 1/(2c)$. (C) Finally, if $a > 1 > b$, $d(k, j)$ does not exceed unity as j moves towards i up to the closer intersection p of the line segment i–j and the unit sphere centred at k. As shown in the inset of Figure 6, the maximal permissible distance $d(j, p)$ is greater than $(1 + c^2 - a^2)/(2c) \geq 1/(2c)$ in this case. This completes the proof of the lemma.

We now show that any biconnected configuration can reach a fully connected configuration after a *finite* number of iterations of the above steps. Suppose this is not true, then there exists a pair $\{i, j\}$ that endures an infinite number of iterations of the above steps. However, this is impossible. Let d_0 be the initial distance $d(i, j)$. Each step applied to $\{i, j\}$, whenever it emerges as the longest pair, reduces the distance $d(i, j)$ by at least $1/[2d(i, j)] \geq 1/(2d_0)$. So, the distance $d(i, j)$ is less than unity after at most $2d_0^2$ steps. At this point, because all other pairs have shorter distances, the configuration is fully connected. This completes the proof.

The above argument shows that it is unnecessary to sample non-biconnected configurations just for ergodicity. It may be tempting to further exclude configurations of zero star contents. This, however, leads to non-ergodic sampling for hard-sphere fluids. For example, one cannot, by a single-point move, transform a configuration of the fully connected diagram ⊠ to that of the ring diagram □, without passing a configuration of ◪ or ◧, whose star content is zero.

Appendix 8. Biconnectivity after a vertex addition

Here, we show that the diagram H obtained from adding a new vertex v to a biconnected diagram G is biconnected if and only if v is adjacent to two distinct vertices in G.

First, if H is biconnected, then v is adjacent to at least two vertices. Otherwise, suppose v is adjacent to only one vertex u in G, then removing u disconnects H, which is a contradiction. Conversely, assuming v is adjacent to two vertices u and w in G, we will prove H is connected and has no articulation point. Since G is connected, $H = G \cup \{v\}$ is connected and v is not an articulation point. Further, since G is biconnected, removing any vertex z from G yields a connected subgraph $G' = G\backslash\{z\}$. It follows that $H' = H\backslash\{z\} = G' \cup \{v\}$ is also connected, for v is adjacent to a vertex u in G other than z, and $u \in G'$. Thus, no vertex in G is an articulation point in H, and H is biconnected.

Appendix 9. Acceptance ratios in the grand ensembles

Here we give explicit expressions of various acceptance ratios in the grand ensemble. In Algorithm G1, the acceptance ratio of the $n \to n + 1$ transition is

$$\langle p_{n\to n+1}\rangle = \int \mathrm{bc}(\mathbf{r}^{n+1})\frac{\mathrm{bc}(\mathbf{r}^n)d\mathbf{r}^n}{Z_n}\frac{\Theta(R - r_{i,n+1})d\mathbf{r}_{n+1}}{R^D V_D}, \quad \text{(I1)}$$

where $\Theta(x)$ is the unit step function (which is 1.0 for a positive x, or 0 otherwise), $\mathrm{bc}(\mathbf{r}^n)/Z_n$ is the probability density of the

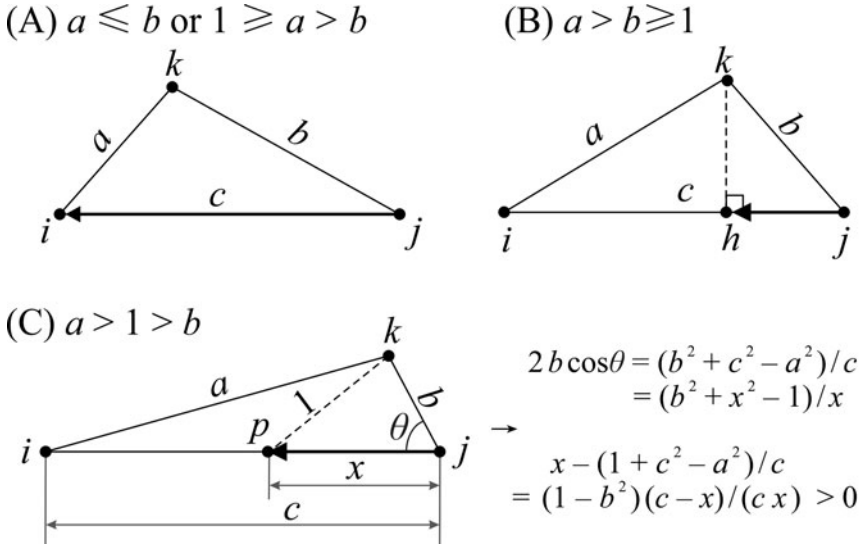

Figure 6. Three cases in proving the lemma for sampling ergodicity: the distance c between the farthest pair of vertices can always be reduced by at least $1/(2c)$ in a single MC step without loss of the biconnectivity of the configuration.

biconnected configuration $\mathbf{r}^n$, and $\Theta(R - r_{i,n+1})/(R^D V_D)$ is that of the new point $\mathbf{r}_{n+1}$. The acceptance ratio of the $n+1 \to n$ transition is

$$\langle p_{n+1\to n}\rangle = \int \Theta(R - r_{i,j})\mathrm{bc}(\mathbf{r}^{n+1}_{-\{i\}})\frac{\mathrm{bc}(\mathbf{r}^{n+1})d\mathbf{r}^{n+1}}{Z_{n+1}}, \qquad \text{(I2)}$$

where $\mathbf{r}^{n+1}_{-\{i\}}$ means $\{\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_{n+1}\}$. From Equations (I1) and (I2), we have Equation (14). Detailed balance is satisfied because the integrand in Equation (I1) is proportional to that in Equation (I2).

In Algorithm G2, the acceptance ratio for the point-removal move is

$$\begin{aligned}\langle p_{n+1\to n}\rangle &= \int \min\left\{1, \frac{N_p(\mathbf{r}^{n+1})}{\bar{N}_p}\right\} \\ &\quad\times \left[\frac{\sum_{i\neq j}\Theta(R - r_{i,j})\mathrm{bc}(\mathbf{r}^{n+1}_{-\{i\}})}{N_p(\mathbf{r}^{n+1})}\right]\frac{\mathrm{bc}(\mathbf{r}^{n+1})d\mathbf{r}^{n+1}}{Z_{n+1}} \\ &= n(n+1)\int \min\left\{1, \frac{N_p(\mathbf{r}^{n+1})}{\bar{N}_p}\right\} \\ &\quad\times \frac{\Theta(R - r_{i,j})\mathrm{bc}(\mathbf{r}^{n+1}_{-\{i\}})}{N_p(\mathbf{r}^{n+1})}\frac{\mathrm{bc}(\mathbf{r}^{n+1})d\mathbf{r}^{n+1}}{Z_{n+1}}, \qquad \text{(I3)}\end{aligned}$$

where the square bracket is equal to 1 for it represents the sum of the selection probabilities of the pairs considered for removal. Since the points are identical, the integral of all ordered pairs (i, j) yields the same result, and the sum is equivalent to the factor $n(n+1)$. The acceptance ratio for the point-addition move is

$$\begin{aligned}\langle p_{n\to n+1}\rangle &= \int \min\left\{1, \frac{\bar{N}_p}{N_p(\mathbf{r}^{n+1})}\right\}\mathrm{bc}(\mathbf{r}^{n+1}) \\ &\quad\times \frac{\mathrm{bc}(\mathbf{r}^n)d\mathbf{r}^n}{Z_n}\frac{\Theta(R - r_{i,n+1})d\mathbf{r}_{n+1}}{R^D V_D}. \qquad \text{(I4)}\end{aligned}$$

From Equations (I3) and (I4), we have Equation (16).

For Algorithm G3, the acceptance ratios of type I transitions are

$$\left\langle p^{0\to 1}\right\rangle = \int \min\{1, 1/\zeta\}\mathrm{bc}(\mathbf{r}^{n+1})\frac{\mathrm{bc}(\mathbf{r}^n)d\mathbf{r}^n}{Z^0}\frac{\Theta(R^1 - r_{i,n+1})d\mathbf{r}_{n+1}}{(R^1)^D V_D} \qquad \text{(I5)}$$

and

$$\begin{aligned}\left\langle p^{1\to 0}\right\rangle &= \int \min\{1, \zeta\}\mathrm{bc}\left(\mathbf{r}^{n+1}_{-\{i\}}\right) \\ &\quad\times \frac{\Theta(R^1 - r_{i,j})\mathrm{c}(\mathbf{r}^{n+1}_{-\{i,j\}})\mathrm{bc}\left(\mathbf{r}^{n+1}\right)d\mathbf{r}^{n+1}}{Z^1}. \qquad \text{(I6)}\end{aligned}$$

From Equations (I5) and (I6), we have Equation (18). Note that the condition $\mathrm{bc}(\mathbf{r}^{n+1}_{-\{i\}}) = 1$ implies $c(\mathbf{r}^{n+1}_{-\{i,j\}}) = 1$, so the latter can be omitted from Equation (I6).

For the transitions between the last intermediate ensemble ($m = M$) and the regular ensemble ($m = M+1$) of $n+1$ points, we have

$$\begin{aligned}\left\langle p^{M\to M+1}\right\rangle &= \int \min\left\{1, \frac{\overline{N}_f}{N_f(\mathbf{r}^{n+1})}\right\} \\ &\quad\times \frac{\mathrm{bc}(\mathbf{r}^{n+1})\mathrm{c}(\mathbf{r}^{n+1}_{-\{i,j\}})\Theta(1 - r_{i,j})d\mathbf{r}^{n+1}}{Z^M}\end{aligned}$$

and

$$\begin{aligned}\left\langle p^{M+1\to M}\right\rangle &= \frac{n(n+1)}{2}\int c(\mathbf{r}^{n+1}_{-\{i,j\}})\min\left\{1, \frac{N_f(\mathbf{r}^{n+1})}{\overline{N}_f}\right\} \\ &\quad\times \frac{\Theta(1 - r_{i,j})}{N_f(\mathbf{r}^{n+1})}\frac{\mathrm{bc}(\mathbf{r}^{n+1})d\mathbf{r}^{n+1}}{Z^{M+1}},\end{aligned}$$

where $n(n+1)/2$ is the number of pairs in the $(n+1)$-point configuration. Equation (19) follows from the above two equations.

For a pair of intermediate ensembles of $n+1$ points, we have

$$\begin{aligned}\left\langle p^{m\to m'}\right\rangle &= \int \Theta(R^{m'} - r'_{i,j})\mathrm{bc}(\mathbf{r}'^{n+1}) \\ &\quad\times \frac{\Theta(R^m - r_{i,j})\mathrm{c}(\mathrm{r}^{n+1}_{-\{i\}})\mathrm{bc}(\mathbf{r}^{n+1})d\mathbf{r}^{n+1}}{Z^m}\end{aligned}$$

and

$$\begin{aligned}\left\langle p^{m'\to m}\right\rangle &= \int \Theta(R^m - r_{i,j})\mathrm{bc}(\mathbf{r}^{n+1}) \\ &\quad\times \frac{\Theta(R^{m'} - r'_{i,j})\mathrm{c}(\mathbf{r}'^{n+1}_{-\{i\}})\mathrm{bc}(\mathbf{r}'^{n+1})d\mathbf{r}'^{n+1}}{Z^{m'}}.\end{aligned}$$

Using Equation (17), we have $\mathbf{r}^{n+1}_{-\{i\}} = \mathbf{r}'^{n+1}_{-\{i\}}$ and $d\mathbf{r}'^{n+1} = s^D d\mathbf{r}^{n+1}$, and Equation (20) follows.

Appendix 10. Asymptotic limit of the ring integral

Here we evaluate the large n limit of the ring integral in Equation (8). In this limit, we can apply the method of steepest descent and focus on the global maximum of $[J_\alpha(k)/k^\alpha]^n$ at $k = 0$, where $\alpha = D/2$ [20]. From the power series of the Bessel function, we have

$$\begin{aligned}\frac{J_\alpha(k)}{k^\alpha} &= \frac{1}{2^\alpha\Gamma(\alpha+1)}\left[1 - \frac{k^2}{4(\alpha+1)} + \cdots\right] \\ &\approx \frac{1}{2^\alpha\Gamma(\alpha+1)}\exp\left[-\frac{k^2}{4(\alpha+1)}\right].\end{aligned}$$

The integral in Equation (8) then yields a gamma function, and [20]

$$\frac{R_n}{B_2^{n-1}} \approx \frac{2^{n-2}(1 - 1/n)}{\Gamma(\alpha+1)}\left(\frac{1+\alpha}{n}\right)^\alpha.$$

In the limit $n \to \infty$, we get $\lim_{n\to\infty}\left(R_n/B_2^{n-1}\right)/\left(R_{n+1}/B_2^n\right) = 1/2$ [20].

Appendix 11. Equations of state under the Percus–Yevick closure

Under the assumption that the indirect correlation function $c(r)$ vanishes beyond the repulsive core $r > 1$ (the PYclosure), the radial distribution function $g(r)$, and hence the virial series, for hard-sphere fluids can be solved analytically in odd dimensions [39–41]. Here we follow Leutheusser's solution [41] for an arbitrary odd dimension $D = 2k+1$, which is reduced to a set of

algebraic equations for the variables Q_n ($n = 0, \ldots, k$) [41]:

$$\begin{aligned}
&-k!2^k Q_k + \lambda \sum_{n=0}^{k} (-1)^n \frac{Q_n}{k+n+1} - (-1)^k = 0,\\
&-(k-1)!2^{k-1} Q_{k-1} + \lambda \sum_{n=0}^{k} (-1)^n \frac{Q_n}{k+n+2} - (-1)^k = 0,\\
&\frac{1}{2}\lambda(-1)^{n+1}[Q^{(n)}(0)]^2 - \lambda \sum_{m=0}^{n-1} (-1)^m Q^{(m)}(0)\, Q^{(2n-m)}(0)\\
&- Q^{(2n+1)}(0) = 0,
\end{aligned} \tag{K1}$$

where $\lambda = (2\pi)^k \rho = (2k+1)!!\rho V_D/2$, and $Q^{(m)}(0)$ is the mth derivative of $Q(r) = \sum_{n=0}^{k} Q_n (r-1)^{n+k}$ evaluated at $r = 0$. The last quadratic equation applies to $0 \le n \le k-2$.

We can reduce Equations (K1) to an implicit polynomial equation $f^{(n)}(Q_n, \rho) = 0$ for each Q_n. Among these solutions, the ones for Q_0 and Q_k directly give the virial and compressibility equations of state, respectively. In the former case, we have

$$\frac{\beta p}{\rho} = 1 + \frac{\rho V_D}{2} g(1^+) = 1 + \frac{\rho V_D}{2}(-1)^{k+1} k! Q_0. \tag{K2}$$

In the latter case, we have

$$\frac{\partial(\beta p)}{\partial \rho} = \frac{\beta}{\rho \kappa_T} = \left|\tilde{Q}(q=0)\right|^2 = (2^k k! Q_k)^2, \tag{K3}$$

where $\tilde{Q}(q) = 1 - \lambda \int_0^1 dr e^{iqr} Q(r)$.

The radius of convergence can also be obtained analytically. Recall that in using the Wiener–Hopf decomposition [40,41]: $1 - \rho\tilde{c}(q) = \tilde{Q}(q)\tilde{Q}(-q)$, where $\tilde{c}(q) = \int d\mathbf{r} e^{i\mathbf{q}\cdot\mathbf{r}} c(r)$, we require a non-vanishing $\tilde{Q}(q)$. This condition holds until ρ moves across a critical negative value ρ_c, which can be found from the case of $q = 0$ as

$$Q_k = \tilde{Q}(0)/[(-)^{k+1} 2^k k!] = 0. \tag{K4}$$

The compressibility κ_T also diverges at this point. Equations (K1) and (K4) yield a polynomial equation $P(\rho_c) = 0$ of ρ_c, which can be readily found from computing the Gröbner basis [57] of the polynomials on the left side of Equations (K1) with Q_k set to 0. The polynomial $P(\rho_c)$ for D up to 19 (which is of degree $2^{(D-3)/2}$) was obtained using MATHEMATICA® [32] and MAGMA [58], and the ρ_c was identified as the negative root of the smallest magnitude (Table 4). The values for $D \le 13$ agreed with the previous literature values [41–43]. In the actual calculation, we used the packing fraction $\eta = \rho V_D/2^D$ instead of the density ρ, for the former yields polynomials of rational coefficients.

Additionally, the virial-route (or compressibility-route) equation of state can also be obtained from the Gröbner basis constructed from Equations (K1), with the change of variable from Q_0 to $g \equiv g(1^+)$ [or from Q_k to $X \equiv \sqrt{\beta/(\rho\kappa_T)}$] according to Equation (K2) (or Equation (K3)). As the final equation of state $P^{(v)}(g, \eta)$ or $P^{(c)}(X, \eta)$ involves two variables, we computed the Gröbner basis by polynomial interpolation, which was faster in our implementation. Given a value η_i ($\eta_i \ne 0,\ 1$), we replace η by η_i in Equations (K1) and construct the specialised Gröbner basis. The resulting polynomial is a multiple of $P^{(v)}(g, \eta_i)$ or $P^{(c)}(X, \eta_i)$. We normalise it such that the leading term is $(1-\eta_i)^{2^k} \eta_i^{2^{k-1}-1} g^{2^{k-1}}$ or $(1-\eta_i)^{2^k} X^{2^{k-1}}$, respectively. We repeat the procedure for $3 \cdot 2^{k-1}$ or $2^k + 1$ different integer values of η_i, and interpolate the normalised polynomials $P^{(v)}(g, \eta_i)$ or $P^{(c)}(X, \eta_i)$ along η to get the respective equation of state $P^{(v)}(g, \eta)$ or $P^{(c)}(X, \eta)$. Note that $P^{(c)}(X = 0, \eta)$ must be a multiple of the polynomial for the radius of convergence obtained earlier. The virial and compressibility equations of state were both computed to $D = 19$. The time needed to construct the Gröbner bases grows exponentially with D, so is the complexity of the equations of state. For example, the virial-route equation of $D = 19$ requires about 180 megabytes to write down, while that of $D = 11$ only requires about 28 kilobytes.

Once we have the equations of state, the virial coefficients from the two routes are computed as follows. Using the compressibility-route as an example, we approximate X by a truncated series $X^{(m)}(\eta) = \sum_{l=0}^{m} X_l \eta^l$, and increase the order m gradually from zero. In each step, we plug the approximant into the equation of state $P^{(c)}[X^{(m)}(\eta), \eta] = 0$, and determine X_m as the value that annihilates the coefficient before η^m in the series expansion of $P^{(c)}$. The first $m+1$ virial coefficients can then be obtained by plugging the truncated series into Equation (K3). The third virial coefficients from both routes must be the same, which furnishes a simple test of validity. The virial series, however, cannot replace the equation of state, for the latter is valid even for an $\eta > |\eta_c|$, where the former diverges.