# Cluster perturbation theory for the self-assembly of associating fluids into complex structures

Bennett D. Marshall

*ExxonMobil Research and Engineering, 3225 Gallows Road, Fairfax, Virginia 22037*

Wertheim's two-density thermodynamic perturbation theory (TPT) has proven to be an indispensable statistical mechanical tool in the description of associating fluids with a single association site. TPT was developed to enforce the monovalence of the hydrogen bond and only recently has been extended to account for divalent association sites. It has been shown through experiment and molecular simulation that certain one-site associating fluids can self-assemble into complex extended supramolecular structures as a result of multiple bonding of association sites. In this paper we reorganize TPT into a form that is more easily applied to complex associated structures. The derived theory is general to all possible self-assemble structures. We obtain the free energy and bonding fractions in a general way in terms of single-cluster partition functions and averages. The new formalism removes any reference to graph theory allowing for the conceptually straightforward application of the two-density formalism to complex self-assembled structures.

## I. INTRODUCTION

Associating fluids contain molecules (or colloids) that interact with strong, short-ranged, and highly directional attractive forces. The most important example of an association interaction is the hydrogen bond; however, more recently, patchy colloids [1–5] have become an important application of association interactions to the self-assembly of soft-matter systems. From a theoretical point of view, association attractions are difficult to model due to the fact that the short-ranged, directional nature of the attraction results in bond saturation. That is the association interaction has a limited valence. Typically, hydrogen-bonding fluids are monovalent; however, the valence of patchy colloid fluids can be tuned by patch size, resulting in an enormous array [6] of possible associated structures.

The formally exact solution for the degree of association in dimerizing (monovalent association site) fluids has been obtained by Anderson [7,8], Chandler and Pratt [9], and Wertheim [10], through very different cluster expansions. When applied as a perturbation theory, where the pair correlation function between associating molecules is taken to be that of the reference fluid, the theory has been shown to be highly accurate [11,12]. It has now been 30 years since Wertheim [10] published his two-density cluster expansion for one-site associating fluids, and only very recently have researchers applied the two-density formalism to one-site associating fluids, which self-assemble into structures [13–16] more complex than simple dimers through multiple bonding of a single association site. While a significant theoretical advance for fluids with a single (and multiple [17–19]) association site, these approaches only account for a limited class of new associated structures such as linear chains of doubly bonded sites [13,15], small rings [14,20], and starlike clusters [16]. However, it has been shown through experiment and molecular simulations [6,21] that one-site associating fluids can self-assemble into a wide range of large and complex supramolecular structures such as lamellae, wormlike strings, micelles, and vesicles [22].

As designed, Wertheim's two-density formalism is most simply applied to the case of a dimerizing fluid. To allow for the application of Wertheim's perturbation theory to one-site fluids which self-assemble into complex structures, it will

prove valuable to rewrite the theory in a form more amiable to application to these types of systems. This will be the subject of this paper. We reorganize Wertheim's theory into a form that shows a clear dependence on the classes of associated clusters that occur. We call this form of the theory "cluster perturbation theory." This new form of the theory is generally applicable to any one-site system and will provide a convenient starting place for the future development of theories for the association of one-site molecules and colloids into specific complex structures.

## II. CLUSTER PERTURBATION THEORY

In this section we reorganize Wertheim's thermodynamic perturbation theory for one-site associating fluids into a form that is convenient for application to associated clusters which contain multibody effects not accounted for in first-order perturbation theory. We begin with a definition of the pair potential,

$$\phi(12) = \phi_{\text{ref}}(12) + \phi_{\text{AA}}(12), \quad (1)$$

where $\phi_{\text{ref}}(12)$ is the reference system pair potential, which contains both short-range repulsions and nonassociation attractions. The association pair potential $\phi_{\text{AA}}(12)$ represents the highly directional, short-ranged association contribution to the pair potential. In total, 5 degrees of freedom describe any single molecule, three translational coordinates represented by the vector $\vec{r}_1$, and two orientation angles represented by $\Omega_1$. These 5 degrees of freedom are represented as $1 = \{\vec{r}_1, \Omega_1\}$. For the potential given in Eq. (1) the Mayer function is decomposed as

$$f(12) = \exp(-\phi(12)/k_B T) - 1 = f_{\text{ref}}(12) + F_{\text{AA}}(12), \quad (2)$$

where

$$F_{\text{AA}}(12) = e_{\text{ref}}(12) f_{\text{AA}}(12)$$
$$e_{\text{ref}}(12) = \exp(-\phi_{\text{ref}}(12)/k_B T) = 1 + f_{\text{ref}}(12)$$
$$f_{\text{AA}}(12) = \exp(-\phi_{\text{AA}}(12)/k_B T) - 1. \quad (3)$$

Using this definition of the pair potential Wertheim developed an exact cluster expansion for one-site associating fluids. Instead of using the density expansion of the pair correlation

function $g(12)$ or Helmholtz free energy $A$, Wertheim uses the fugacity expansion of $\ln \Xi$, where $\Xi$ is the grand partition function, as the starting point. Building on the ideas of Lockett [23], Wertheim then regroups the expansion such that individual graphs are composed of $s$-mer graphs. An $s$-mer represents a cluster of points that are connected by paths of $F_{AA}$ bonds, each pair of points in an $s$-mer, which are not directly connected by a $F_{AA}$ bond, receives an $e_{ref}(12)$ bond. This regrouping serves to include the geometry of association with the $e_{ref}(12)$ bonds enforcing the limited valence of the association interaction. In the $s$-mer representation, graphs which include unphysical core overlap are identically zero. That is, if the association site is singly bondable all graphs composed of $s$-mers of size $s > 2$ immediately vanish due to steric hindrance.

This regrouping of the fugacity expansion allows for the easy incorporation of steric effects. Wertheim uses the idea of multiple densities, splitting the total density of the fluid as

$$\rho(1) = \rho_o(1) + \rho_b(1), \tag{4}$$

where $\rho_o(1)$ is the density of monomers (molecules not bonded) and $\rho_b(1)$ is the density of molecules that are bonded. The density $\rho_o(1)$ is composed of all graphs in $\rho(1)$ that do not have an incident $F_{AA}$ bond, and $\rho_b(1)$ contains all graphs that have one or more incident $F_{AA}$ bonds. Performing a topological reduction from fugacity graphs to graphs that contain $\rho_o(1)$ and $\rho(1)$ field points, allowed Wertheim to arrive at the following exact free energy:

$$\frac{A}{k_B T} = \int (\rho(1) \ln(\rho_o(1)\Lambda^3) - \rho_o(1))d(1) - c^{(o)}, \tag{5}$$

where the graph sum $c^{(o)}$ is given as

$$c^{(o)} = \begin{Bmatrix} \text{sum of all irreducible graphs consisting of monomer points carrying factors} \\ \text{of } \rho, s\text{-mer graphs with } s \geqslant 2 \text{ and every point carrying a factor of } \rho_o, \text{and} \\ f_{ref}\text{-bonds between some sets of points in distinct } s\text{-mers.} \end{Bmatrix} \tag{6}$$

To develop a form of Wertheim's TPT that is suitable for application to associated clusters that contain multiply bonded sites, it is convenient to recast the theory into a form that illustrates the classes of associated clusters. In TPT, all contributions to $c^{(o)}$ that contain more than one associated cluster are neglected. This allows the remaining diagrams to be summed in terms of reference system correlation functions. Since each contribution to $\Delta c^{(o)} = c^{(o)} - c_{ref}^{(o)}$ contains a single type of associated cluster ($c_{ref}^{(o)}$ contains graphs with no association bonds), it makes sense to write the sum $\Delta c^{(o)}$ as a sum over individual cluster types,

$$\Delta c^{(o)} = \sum_v^{\text{all clusters}} \Delta c_v, \tag{7}$$

where $\Delta c_v$ gives the sum of all diagrams that contain a single associated cluster type $v$ interacting with the reference fluid. For a homogeneous fluid this infinite sum of diagrams for each cluster type can be summed in terms of reference system correlation functions as

$$\frac{\Delta c_v}{V} = \frac{1}{S(v)} \frac{\rho_o^{n(v)}}{\Omega^{n(v)-1}} \int g_{ref}(1 \cdots n(v)) \\ \times \prod_{\substack{\text{bonded} \\ \text{pairs}}} f_{AA}(lk)d(2) \cdots d(n(v)), \tag{8}$$

where $S(v)$ is the cluster symmetry number, $n(v)$ is the number of molecules in the associated cluster, and $g_{ref}$ is the multibody correlation function.

An alternative representation of Eqs. (7) and (8) can be obtained by introduction of single-cluster averages and partition functions. We begin by elimination of the multibody correlation functions $g_{ref}(1 \cdots s)$ in favor of the cavity correlation functions $y_{ref}(1 \cdots s)$ defined as

$$y_{ref}(1 \cdots s) \prod_{\text{labeled pairs}} e_{ref}(lk) = g_{ref}(1 \cdots s). \tag{9}$$

The probability that the isolated cluster $v$ has a configuration $(1 \ldots n(v))$ is given by

$$P_v(1 \cdots n(v)) = \frac{\prod_{\text{all pairs}} e_{ref}(lk) \prod_{\substack{\text{bonded} \\ \text{pairs}}} f_{AA}(12)}{Z_v}, \tag{10}$$

where the single cluster partition function $Z_v$ is given by

$$Z_v = \int \prod_{\text{all pairs}} e_{ref}(lk) \prod_{\substack{\text{bonded} \\ \text{pairs}}} f_{AA}(12)d(2) \cdots d(n(v)). \tag{11}$$

Equations (9)–(11) now allow Eq. (8) to be recast as

$$\frac{\Delta c_v}{V} = \frac{1}{S(v)} \frac{\rho_o^{n(v)}}{\Omega^{n(v)-1}} \langle y_{ref}(1 \cdots n(v)) \rangle_v Z_v. \tag{12}$$

The braces $\langle \rangle_v$ represent a single cluster average over the single cluster probability Eq. (10).

Now using Eqs. (7) and (12) the Helmholtz free energy is minimized with respect to monomer density giving the following relation for the density of molecules bonded:

$$\rho_b = \sum_v^{\text{all clusters}} \rho_b^{(v)}, \tag{13}$$

where $\rho_b^{(v)}$ is the density of molecules bonded in clusters of type $v$, given by the relation

$$\rho_b^{(v)} = \frac{n(v)}{S(v)} \frac{\rho_o^{n(v)}}{\Omega^{n(v)-1}} \langle y_{ref}(1 \cdots n(v)) \rangle_v Z_v. \tag{14}$$

The only unknown in Eq. (13) is the monomer densities, which are obtained self consistently through Eqs. (4), (13), and (14). Comparing Eqs. (12) and (14) we see

$$\frac{\Delta c_v}{V} = \frac{\rho_b^{(v)}}{n(v)}. \tag{15}$$

Taking the homogeneous limit of Eq. (5) and employing Eq. (15) we obtain

$$\frac{A - A_{\text{ref}}}{N k_b T} = \ln(1 - X_b) + X_b - \sum_{v}^{\text{all clusters}} \frac{X_b^{(v)}}{n(v)}, \qquad (16)$$

where $N$ is the number of molecules, $X_b^{(v)} = \rho_b^{(v)}/\rho$ is the fraction of molecules bonded in cluster type $v$, and $X_b = \rho_b/\rho$ is the fraction of molecules bonded in any cluster,

$$X_b = \sum_{v}^{\text{all clusters}} X_b^{(v)}. \qquad (17)$$

Further, if we introduce the fraction of molecules in associated clusters of size $k$,

$$X_c^{(k)} = \sum_{n(v)=k} X_b^{(v)}, \qquad (18)$$

where

$$X_b = \sum_{k=2}^{\infty} X_c^{(k)}, \qquad (19)$$

the free energy can be written as

$$\frac{A - A_{\text{ref}}}{N k_b T} = \ln(1 - X_b) + X_b - \sum_{k=2}^{\infty} \frac{X_c^{(k)}}{k}. \qquad (20)$$

The results presented in this section provide a general reorganization of TPT, which will allow for the conceptually straightforward extension for systems that go beyond simple dimerization. This general solution of TPT can be applied by researchers without a background in graph theory or Wertheim's multidensity approach.

## III. APPROXIMATE CLUSTER DENSITIES FOR PATCHY HARD SPHERES

The results in Sec. II give a general reorganization of TPT into a form that eliminates any reference to graph theory. In general, the averages $\langle y_{\text{ref}}(1 \cdots n(v)) \rangle_v$ and cluster partition functions $Z_v$ will be temperature dependent. The theory can be significantly simplified if we consider a hard-sphere reference fluid with conical square well association sites. These patchy hard spheres have been used as primitive models for hydrogen-bonding fluids for many decades. The patchy hard-sphere potential, first introduced by Bol [24] and later used by Chapman *et al.* [12,25], considers association as a square well interaction, which depends on the position and orientation of each molecule. Kern and Frenkel [26] later realized that this potential could describe the interaction between "patchy" colloids. For these conical sites the association potential is given by

$$\phi_{\text{AA}}(12) = -\varepsilon_{\text{AA}} O_{\text{AA}}(12)$$

$$O_{\text{AA}}(12) = \begin{cases} 1, & r_{12} \leqslant r_c; \ \theta_{A1} \leqslant \theta_c; \ \theta_{A2} \leqslant \theta_c \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\text{AA}}(12) = (\exp(\varepsilon_{\text{AA}}) - 1) O_{\text{AA}}(12) = f_{\text{AA}} O_{\text{AA}}(12) \quad (21)$$

where $r_c$ is the maximum separation between two colloids for which association can occur, $\theta_{A1}$ is the angle between $\vec{r}_{12}$ and
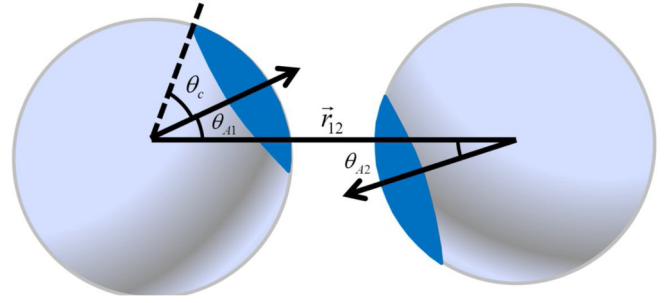


FIG. 1. (Color online) Association parameters for conical association site.

the orientation vector passing through the center of the patch on colloid 1, and $\theta_c$ is the critical angle beyond which association cannot occur. Equation (21) states that if the spheres are close enough, $r_{12} < r_c$, and both are oriented correctly, $\theta_{A1} < \theta_c$ and $\theta_{A2} < \theta_c$, then an association bond is formed and the energy of the system is decreased by $\varepsilon_{\text{AA}}$. Figure 1 gives an illustration of two single-site spheres interacting with this potential. The size of the patch is dictated by the critical angle $\theta_c$, which defines the solid angle to be $2\pi (1 - \cos \theta_c)$. The patch size determines the maximum number of other spheres to which the patch can bond. Specifically considering a hard-sphere reference fluid with association occurring at hard-sphere contact $r_c = d$, it is possible for a patch to associate at most once for $0° \leqslant \theta_c < 30°$, twice for $30° \leqslant \theta_c < 35.3°$, three times for $35.3° \leqslant \theta_c < 45°$, and four times for $45° \leqslant \theta_c < 58.3°$ [13].

The advantage of employing this patchy hard-sphere model is twofold. This choice of potential renders the cluster averages $\langle y_{\text{ref}}(1 \cdots n(v)) \rangle_v$ independent of temperature. For this patchy hard-sphere model it is beneficial to define a new temperature independent cluster partition function,

$$\tilde{Z}_v = \int \prod_{\text{all pairs}} e_{\text{HS}}(lk) \prod_{\substack{\text{bonded} \\ \text{pairs}}} O_{\text{AA}}(12) d(2) \cdots d(n(v)). \tag{22}$$

Equation (22) is simply the total number of associated states of isolated cluster $v$. For the patchy hard-sphere case the cluster graph sums and densities can now be rewritten as

$$\frac{\Delta c_v}{V} = \frac{\rho_b^{(v)}}{n(v)} = \frac{(f_{\text{AA}})^{n(f_{\text{AA}})}}{S(v)} \frac{\rho_o^{n(v)}}{\Omega^{n(v)-1}} \langle y_{\text{HS}}(1 \cdots n(v)) \rangle_v \tilde{Z}_v, \tag{23}$$

where $f_{\text{AA}} = \exp(\varepsilon_{\text{AA}}/k_B T) - 1$ and $n(f_{\text{AA}})$ is the number of association bonds in the cluster. Of course, it may be desirable to include spherically symmetric attractions (Lennard-Jones, square well, etc.) in addition to association interactions. These are typically included in the high-temperature expansion [27,28].

The challenge in applying TPT to a given system has now been reduced to determining the relevant types of associated clusters, evaluating the cluster averages $\langle y_{\text{HS}}(1 \cdots n(v)) \rangle_v$ and cluster partition functions $\tilde{Z}_v$. For the case of a dimerizing fluid, these quantities are easily obtained as

$$\langle y_{\text{HS}}(12) \rangle_{\text{dimer}} = \frac{4\pi \int_d^{r_c} y_{\text{HS}}(r) r^2 dr}{4\pi \int_d^{r_c} r^2 dr} = \frac{\xi}{v_b}$$

$$\tilde{Z}_{\text{dimer}} = v_b \kappa_{\text{AA}} \Omega \tag{24}$$

where $v_b$ is the volume of a spherical shell of thickness $r_c - d$, $\Omega = 4\pi$ is the total number of orientations, and $\kappa_{AA}$ is the probability that two spheres are oriented correctly for association to occur

$$\kappa_{AA} = (1 - \cos\theta_c)^2/4. \qquad (25)$$

When considering clusters larger than dimers $n(v) > 2$, evaluation of the single-cluster averages and partition functions become more complex. A particularly simple approximation, which will prove convenient, is to approximate the cluster averages as

$$\langle y_{HS}(1 \cdots n(v)) \rangle_v \approx \langle y_{HS}(12) \rangle_{dimer}^{n(f_{AA})} = \left(\frac{\xi}{v_b}\right)^{n(f_{AA})}. \qquad (26)$$

In Eq. (26) we have approximated the multibody cavity correlation functions by a superposition in which each pair of spheres that share an association bond also receive a pair cavity correlation function. This approximation treats each association bond independently in regards to the cluster averages, with multibody ($>2$) effects completely described by the cluster partition functions. This is analogous to density functional theories for polyatomic molecules, which treat chain connectivity at the ideal level and add density dependence in the excess terms [29–31]. Here, this approximation will be exact in the low-density limit.

To evaluate the cluster partition functions we employ the mean value theorem [32] to obtain

$$\tilde{Z}_v = \tilde{Z}_{dimer}^{n(v)-1} P_{gen}^{(v)} = (v_b \kappa_{AA} \Omega)^{n(v)-1} P_{gen}^{(v)}, \qquad (27)$$

where $\tilde{Z}_{dimer}^{n(v)-1}$ is the total integration volume and $P_{gen}^{(v)}$ is a generation probability, which is defined by Eq. (27) and evaluated using Monte Carlo integration techniques. Evaluation of $P_{gen}^{(v)}$ for certain cluster classes will be discussed in Sec. IV. Combining these results we obtain the final general form for the cluster densities:

$$\frac{\rho_b^{(v)}}{\rho_o} = \frac{n(v)}{S(v)} \left(\frac{f_{AA}\xi}{v_b}\right)^{n(f_{AA})} (\rho_o v_b \kappa_{AA})^{n(v)-1} P_{gen}^{(v)}. \qquad (28)$$

Equations (16) and (28) provide a general starting point for the development of perturbation theories for one-site associating fluids. Application of the theory follows the following steps:

(1) Identify the relevant types of associated clusters. For each cluster type $v$ the terms $n(v)$, $n(f_{AA})$, and $S(v)$ are obtained by simple inspection.

(2) For each cluster type $v$, the generation probability $P_{gen}^{(v)}$ must be calculated using Monte Carlo integration. For patchy hard spheres, $P_{gen}^{(v)}$ is independent of temperature and density, so it only needs to be computed once for each choice of $\vartheta_c$ and $r_c$. $P_{gen}^{(v)}$ is the workhorse of the theory, encoding the geometric constraints of forming multiple association bonds per association site.

(3) The monomer densities are obtained self-consistently through the relationship $\rho = \rho_o + \sum_v^{all\ clusters} \rho_b^{(v)}$.

(4) Evaluate the Helmholtz free energy (or resulting chemical potentials/pressure) through Eqs. (16) or (20).

In what follows, we demonstrate application of the theory for the case of patchy hard spheres with divalent association sites.
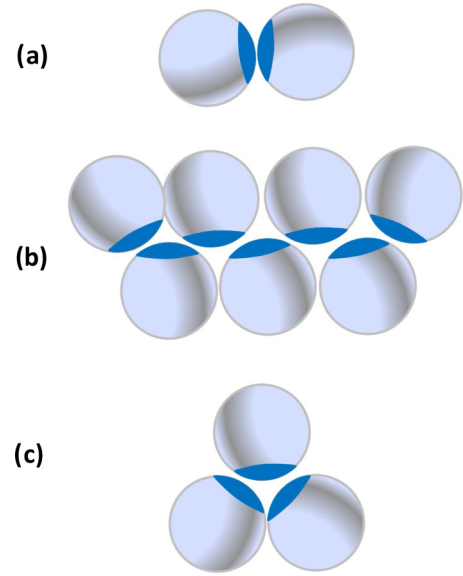


FIG. 2. (Color online) Dominant associated clusters for patchy colloids with a single divalent patch. (a) An example of dimers composed of two spheres with a singly bonded patch, (b) chains of spheres where chain centers of doubly bonded patches, and (c) triatomic rings where all three spheres have doubly bonded patches.

## IV. DIVALENT ASSOCIATION SITES

One of the main assumptions in the development of Wertheim's first-order perturbation theory (TPT1) [33] is that association sites are monovalent. Indeed, the entire two-density formalism of Wertheim is constructed to enforce this condition. For the case of hydrogen bonding, the assumption of singly bondable sites is justified; however, for patchy colloids it has been shown experimentally [3,5] that the number of bonds per patch (association site) is dependent on the patch size. If TPT is to describe these types of systems, the possibility of multiple bonds per association site must be accounted for. For the case of a pure component fluid of divalent single-patch colloids, the dominant types of associated clusters are illustrated in Fig. 2. We follow the recent approach of Marshall [20] who summed over all possible chain lengths, in addition to including the triatomic rings. Our purpose here is to demonstrate how the cluster average perturbation theory allows for the rapid development of theories for associating fluids.

Table I lists the relevant quantities needed to evaluate the cluster densities, Eq. (28). Using the quantities in Table I, the density of spheres bonded in triatomic rings is easily obtained from Eq. (28) as

$$\rho_b^{(ring)} = \frac{1}{2 v_b \kappa_{AA}} (f_{AA} \xi \kappa_{AA} \rho_o)^3 P_{gen}^{(ring)}, \qquad (29)$$

TABLE I. Structural quantities required to evaluate the cluster densities.

| V | $S(v)$ | $n(v)$ | $n(f_{AA})$ |
|---|---|---|---|
| Chains | 2 | $m$ | $m$ - 1 |
| Rings | 6 | 3 | 3 |

where $P_{\text{gen}}^{(\text{ring})}$ is evaluated using Eq. (27) and Monte Carlo integration as

$$P_{\text{gen}}^{(\text{ring})} = \left\{ \begin{array}{l} \text{The probability that if the positions and orientations of two colloids are} \\ \text{generated such that they are positioned and oriented correctly to bond to} \\ \text{a third colloid, that there is no core overlap between the two generated} \\ \text{colloids and that these two generated colloids are oriented and positioned} \\ \text{such that they share an association bond.} \end{array} \right\} \qquad (30)$$

For the densities of spheres bonded in chains of length $m$ we have

$$\frac{\rho_b^{(m)}}{\rho_o} = \frac{m}{2}(\rho_o f_{\text{AA}}\xi\kappa_{\text{AA}})^{m-1} P_{\text{gen}}^{(m)}. \qquad (31)$$

All that remains now is the approximation of $P_{\text{gen}}^{(m)}$. For large chain lengths the rigorous evaluation of this generation probability would be intractable. However, since the association of divalent colloids into chains longer than dimers is dominated by the three-body generation probability $P_{\text{gen}}^{(3)}$, $P_{\text{gen}}^{(m)}$ can be factored in terms of the triatomic generation probability as

$$P_{\text{gen}}^{(m)} = \left( P_{\text{gen}}^{(3)} \right)^{m-2}. \qquad (32)$$

The approximation Eq. (32) is similar to that made by Kalyuzhnyi *et al.* [13] [in their case the $e_{\text{HS}}(r_{k,k+2})$ functions are replaced by hard-sphere reference Mayer functions $f_{\text{HS}}(r_{k,k+2})$], which was shown by explicit comparison to Monte Carlo integration results for the full chain partition function to be accurate. Finally, $P_{\text{gen}}^{(3)}$ is easily evaluated as

$$P_{\text{gen}}^{(3)} = \left\{ \begin{array}{l} \text{The probability that if the positions of two colloids are generated such} \\ \text{that they are correctly positioned to associate with a third colloid, that} \\ \text{there is no core overlap between the two generated colloids.} \end{array} \right\} \qquad (33)$$

Now the (infinite) cluster sum in Eq. (16) can be evaluated as

$$\rho \sum_{v}^{\text{all clusters}} \frac{X_b^{(v)}}{n(v)} = \frac{1}{6v_b\kappa_{\text{AA}}}(f_{\text{AA}}\xi\kappa_{\text{AA}}\rho_o)^3 P_{\text{gen}}^{(\text{ring})} + \frac{1}{2}\frac{\rho_o^2 f_{\text{AA}}\xi\kappa}{1 - f_{\text{AA}}\xi\kappa\rho_o P_{\text{gen}}^{(3)}}. \qquad (34)$$

Similarly, the infinite sum in Eq. (13) can now be evaluated to yield the total density of bonded spheres, which when combined with Eq. (4) gives a closed equation for the monomer density,

$$\rho = \rho_o + \frac{\rho_o^2 f_{\text{AA}}\xi\kappa_{\text{AA}}}{1 - f_{\text{AA}}\xi\kappa_{\text{AA}}\rho_o P_{\text{gen}}^{(3)}} + \frac{1}{2}\rho_o^3 P_{\text{gen}}^{(3)} \left( \frac{f_{\text{AA}}\xi\kappa_{\text{AA}}}{1 - f_{\text{AA}}\xi\kappa_{\text{AA}}\rho_o P_{\text{gen}}^{(3)}} \right)^2 + \frac{1}{2v_b}(\rho_o f_{\text{AA}}\xi)^3 \kappa_{\text{AA}}^2 P_{\text{gen}}^{(\text{ring})}. \qquad (35)$$

Equations (16), (34), and (35) give the theory for single-patch divalent colloids and have been shown [20] to be accurate in comparison to molecular simulation data. As can be seen, the final equations were obtained with relative ease as compared to the approach of Marshall [20], which started with the free energy Eq. (5). Detailed numerical results for the generation probabilities Eqs. (30) and (33) can be found in the original [20] work.

## V. CONCLUSIONS

We have reorganized Wertheim's thermodynamic perturbation theory for one-site associating fluids into a general form that may be applied to fluids that self-assemble into complex structures. The derived theory allows for all possible types of associated clusters. Equations (14) and (16) give the general solutions for the cluster densities and free energy for associating fluids with an arbitrary reference and association pair potential, while Eq. (28) gives an approximate form of the cluster densities for patchy hard spheres. To apply the patchy hard-sphere theory, one needs to identify the possible cluster types, identify their symmetry number, and calculate the single cluster probabilities $P_{\text{gen}}^{(v)}$.

Development of $P_{\text{gen}}^{(v)}$ for complex associated structures is by no means a trivial task. For cases where direct evaluation of $P_{\text{gen}}^{(v)}$ proves prohibitively complex, approximations may be employed as in Sec. IV. Alternatively, since $P_{\text{gen}}^{(v)}$ is independent of density, it should be possible to fit $P_{\text{gen}}^{(v)}$ to simulation data at low density and use these generation probabilities for higher density fluids.

The most significant contribution of this work is the general reorganization of Wertheim's perturbation theory, such that familiarity with graph theory and the complexities of Wertheim's multidensity approach is not required for application of the theory to new systems.

[1] E. Bianchi, R. Blaak, and C. N. Likos, Phys. Chem. Chem. Phys. **13**, 6397 (2011).

[2] Q. Chen, S. C. Bae, and S. Granick, Nature **469**, 381 (2011).

[3] L. Feng, R. Dreyfus, R. Sha, N. C. Seeman, and P. M. Chaikin, Adv. Mater. **25**, 2779 (2013).

[4] A. B. Pawar and I. Kretzschmar, Langmuir **24**, 355 (2008).

[5] Y. Wang, Y. Wang, D. R. Breed, V. N. Manoharan, L. Feng, A. D. Hollingsworth, M. Weck, and D. J. Pine, Nature **491**, 51 (2012).

[6] G. Munaò, Z. Preisler, T. Vissers, F. Smallenburg, and F. Sciortino, Soft Matter **9**, 2652 (2013).

[7] H. C. Andersen, J. Chem. Phys. **59**, 4714 (1973).

[8] H. C. Andersen, J. Chem. Phys. **61**, 4985 (1974).

[9] D. Chandler and L. R. Pratt, J. Chem. Phys. **65**, 2925 (1976).

[10] M. Wertheim, J. Stat. Phys. **35**, 19 (1984).

[11] W. G. Chapman, J. Chem. Phys. **93**, 4299 (1990).

[12] G. Jackson, W. G. Chapman, and K. E. Gubbins, Mol. Phys. **65**, 1 (1988).

[13] Y. Kalyuzhnyi, H. Docherty, and P. Cummings, J. Chem. Phys. **133**, 044502 (2010).

[14] B. D. Marshall, D. Ballal, and W. G. Chapman, J. Chem. Phys. **137**, 104909 (2012).

[15] B. D. Marshall and W. Chapman, Soft Matter **10**, 5168 (2014).

[16] B. D. Marshall and W. G. Chapman, J. Chem. Phys. **139**, 104904 (2013).

[17] Y. Kalyuzhnyi and P. Cummings, J. Chem. Phys. **139**, 104905 (2013).

[18] Y. Kalyuzhnyi, H. Docherty, and P. Cummings, J. Chem. Phys. **135**, 014501 (2011).

[19] Y. Kalyuzhnyi, B. Marshall, W. Chapman, and P. Cummings, J. Chem. Phys. **139**, 044909 (2013).

[20] B. D. Marshall, Ph.D. thesis, Rice University, 2014, http://hdl.handle.net/1911/77223.

[21] S. Jiang, Q. Chen, M. Tripathy, E. Luijten, K. S. Schweizer, and S. Granick, Adv. Mater. **22**, 1060 (2010).

[22] F. Sciortino, A. Giacometti, and G. Pastore, Phys. Rev. Lett. **103**, 237801 (2009).

[23] A. Lockett III, J. Chem. Phys. **72**, 4822 (1980).

[24] W. Bol, Mol. Phys. **45**, 605 (1982).

[25] W. G. Chapman, Ph.D. thesis, Coronell University, Ithaca, NY, 1988.

[26] N. Kern and D. Frenkel, J. Chem. Phys. **118**, 9882 (2003).

[27] A. Gilvillegas, A. Galindo, P. J. Whitehead, S. J. Mills, G. Jackson, and A. N. Burgess, J. Chem. Phys. **106**, 4168 (1997).

[28] J. Gross and G. Sadowski, Ind. Eng. Chem. Res. **41**, 5510 (2002).

[29] D. Cao and J. Wu, J. Chem. Phys. **121**, 4210 (2004).

[30] C. E. Woodward, J. Chem. Phys. **94**, 3183 (1991).

[31] Y. Yu and J. Wu, J. Chem. Phys. **117**, 2368 (2002).

[32] D. Frenkel and B. Smit, *Understanding Molecular Simulation: from Algorithms to Applications* (Academic Press, New York, 2001).

[33] M. S. Wertheim, J. Stat. Phys. **35**, 35 (1984).