Published in final edited form as:

J Chem Phys. 2007 November 21; 127(19): 194506. doi:10.1063/1.2785188.

A site-renormalized molecular fluid theory

Kippi M. Dyer^{a)}, John S. Perkyns^{b)}, and B. Montgomery Pettitt^{c)}

Chemistry Department, University of Houston, Houston, Texas 77204-5003, USA

Abstract

The orientation-dependent pair distribution function for molecular fluids on site-site potentials is expanded in a topological analog of the diagrammatically proper site-site theory of liquids [D. Chandler *et al.*, *Mol. Phys.* **46**, 1335 (1982)]. The resulting functions are then used to diagrammatically renormalize the molecular fluid theory. A result is that the diagrammatically proper interaction site model theory is shown to be a linearized, minimal angular basis set approximation to this site-renormalized molecular theory. This framework is used to propose a new, exact, and proper closure to the diagrammatically proper interaction site model theory. The resulting equation system contains a bridge function expansion in the proper site-site theory. In addition, the construction of the theory is such that the molecular pair distribution function, in full dimensionality, is intrinsic to the theory. Furthermore, the theory is equivalent to the molecular Ornstein-Zernike treatment of site-site molecules in the basis set expansion of Blum and Torruella [*J. Chem. Phys.* **56**, 303 (1971)]. A significant formal result of the theory is the demonstration that certain classes of diagrams which would otherwise be considered improper in the interaction site model formalism are included in the angular expansion of molecular interactions. Numerical results for several apolar homonuclear models and an apolar heteronuclear model are shown to quantitatively improve upon those of reference interaction site model and our recent proper variant with respect to simulation. Significant numerical results are that the various thermodynamic quantities obey the exact symmetries and sum rules within numerical error for the different sites in the heteronuclear case, even for the low order approximation used in this work, and the theory is independent of the so-called auxiliary site problem common to previous site-site theories.

I. Introduction

Integral equation methods for single-center angular expansion models of molecular fluids have been well explored for some time.^{1,2} These methods continue to be extended in application to new single-center models,³ reactive models,⁴ and self-consistent electronic structure calculations.⁵ For molecular potentials defined on atomic sites, however, these more general methods have usually been neglected in favor of the reduced dimensionality of the site-site radial distribution function methods.^{6,7} There have also been efforts to use intermediate, three-dimensional (3D) functions.^{8–10} These methods, especially the reference interaction site model¹¹ (RISM) and the diagrammatically proper interaction site model^{12,13} (PISM) equations, have been applied across many areas of interest in the theory of liquids. The reduced dimensionality of the radial distribution functions makes them much more numerically convenient than the more general angular functions. However, quantitatively predictive results from the ISM equations have proven difficult to refine.^{13–15}

^{a)}Also at Chemistry Department, SUNY-Stony Brook, Stony Brook, NY 11794. Electronic mail: dyer@kitten.chem.uh.edu

^{b)}Electronic mail: jperkyns@uh.edu

^{c)}Electronic mail: pettitt@uh.edu

The efforts to produce quantitatively predictive results for site-site models share certain basic problems. At the most fundamental level, as in the simple fluid case, the closure equation¹⁶ used in a particular level of theory can be viewed as determining the formal exactness of any particular approach. The formal system is further complicated in the 3D and site-site pair functions by the information lost in the angle averaging procedure. This is directly reflected by an increase in the number and types of bridge function terms which must be accounted for explicitly at each level of reduction. In the more general cases, the dimensionality of the full molecular and 3D cases requires a level of numerical complexity not necessary to the radial ISM treatments. In the full molecular case, the angular basis set methods used require a large number of terms for convergence on detailed models, while the 3D case requires a substantial grid and thus significant computer memory access to obtain converged answers.

In recent work,¹⁷ we took advantage of the molecular equations to propose a simple bridge function approximation in the proper interaction site theory. The result was an effective density theory which gave very encouraging results for the site-site pair functions of a simple homonuclear model. In this work, our goal is to address the molecular equations directly, by explicitly expanding the orientation-dependent molecular pair distribution functions in site-based diagrams.

We expand the molecular Ornstein-Zernike (OZ) equation as an n -component mixture in order to simultaneously generate the molecular distribution functions on the various and sometimes symmetrically degenerate site origins. We separate the molecular functions into orientation-dependent functions analogous to the radially symmetric equivalents in the diagrammatically proper site-site theory. We then show that this decomposition allows us to resum or renormalize over angles the site-site molecular potential in a manner similar to the Coulomb renormalization commonly used for simple fluids. We use the renormalized equations to show that the site-site hypernetted-chain (HNC)-like PISM equations¹³ are equivalent to one particular, renormalized linearization of the orientation-dependent functions truncated at a minimal, isotropic basis set approximation. We then generate a new, exact, proper closure which includes the full molecular potential by construction. For diatomic molecules, the resulting equations have a similar complexity and form to the four-point h-bond bridge diagram expansions utilized in simple fluid mixtures.^{18,19} Finally, by including a restricted chain sum²⁰ of cross terms from the molecular (angular-dependent) OZ equation, which are not included in the PISM OZ equations, in combination with the new closure relation, we generate a set of ISM equations that are fully equivalent to the orientation-dependent molecular fluid theory.¹ A significant formal result of this equivalence is that certain subclasses of diagrams which would otherwise be considered diagrammatically improper in the ISM formalism are properly included in the context of the complete angular description of the interactions of molecular fluids.

We give numerical results for several homonuclear and heteronuclear, apolar, Lennard-Jones model diatomic models in comparison to RISM, PISM, and our effective density variant of PISM, all with respect to simulation. We present both the site-site radial distribution function, $g(r)$, and the next, more complicated pair function, $\langle g(r, \Omega_1, \Omega_2) \rangle_{\Omega_2}$, the diatomic analog of the 3D pair distribution function. The results here are for the lowest order, isotropic approximation to the full angular basis expansion. However, even in this simple approximation, the numerical results are a significant improvement over the previous methods. An important result is that the various thermodynamic results measured obey the expected symmetries and sum rules within numerical error for all site pairs. Finally, by setting the interaction potential of one of the sites in a heteronuclear model to zero, we show that this theory does not have the so-called auxiliary site problem that is otherwise a part of the RISM and PISM class of theories.

II. Theory

A. Angular-dependent site-site correlations

We begin by considering the molecular pair distribution,^{16,21,22} $g(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)$, of a site-site model molecular fluid. We use standard notation, where Ω_i refers to the orientation of molecule i , with the additional site-site species labeling of the displacement vector, $\mathbf{r}_{\alpha\gamma}$, where we mean the displacement of molecule 2 from molecule 1 measured between site α on molecule 1 and γ on molecule 2 in the coordinate frame coincident with site α on molecule 1. That is, we choose to simultaneously represent the molecular distribution functions in the n^2 possible equivalent coordinate systems coincident with the site-site pairs between molecules, where there are n different sites on each molecule. In this case, the complete complement of direct correlation functions, $c(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)$, can be generated at once through the mixture form of the molecular OZ equation in Fourier space,

$$\mathbf{h} = \mathbf{c} + \langle c\rho\mathbf{h} \rangle_{\Omega_3}, \quad (1)$$

where $\mathbf{c}$ represents the species labeled matrices whose elements are the Fourier transforms $c(\mathbf{k}_{\alpha\gamma}, \Omega_1, \Omega_2)$, $\mathbf{h}$ represents the total correlation functions defined as $h=g-1$, and ρ is the diagonal matrix with $\rho_{\alpha\alpha}=\rho/n$ for ρ the molecular number density of the system and n the number of sites in the molecule. In a standard convention, the brackets represent the Ω_3 -weighted average over the orientation of molecule 3. For this system, the OZ equation is closed by

$$c(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) = e^{-\beta U(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + t(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + d(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) - t(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)} - 1, \quad (2)$$

where $\beta^{-1}=k_B T$, k_B is the Boltzmann constant, T is the absolute temperature, $U(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)$ is the total intermolecular potential, $t=h-c$ is the indirect correlation function, and $d(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)$ is the bridge function. For the case of a heteronuclear diatomic, the site-site intermolecular potential has a convenient form

$$U(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) = u_{\alpha\gamma}(r_{\alpha\gamma}) + u_{\gamma\gamma}(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1|) + u_{\alpha\alpha}(|\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2|) + u_{\gamma\alpha}(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1 + \mathbf{l}_2|), \quad (3)$$

where $\mathbf{l}_i$ is the intramolecular bond vector between the sites of molecule i , and $u_{\alpha\gamma}(r)$ is the potential between site α on molecule 1 and site γ on molecule 2. Throughout this work on diatomics, our labeling of $\mathbf{r}$ and $\mathbf{l}$ is an assumed reduction of Chandler's labeling convention,⁷ $\mathbf{r}_{\alpha\gamma}=\mathbf{r}_{1\alpha 2\gamma}$ and $\mathbf{l}_i=\mathbf{l}_{i\alpha i\gamma}$, respectively, where $\mathbf{r}_{1\alpha 2\gamma}$ is the vector displacement between site α on molecule 1 and site γ on molecule 2, while $\mathbf{l}_{i\alpha i\gamma}$ is the vector displacement between sites α and γ , both on molecule i . As used here, then, the site label pairs $\alpha\alpha$, $\alpha\gamma$, $\gamma\gamma$ as occur in, e.g., Eq. (3) always refer to distinct site pairs on different molecules, and $\alpha \neq \gamma$ throughout. We use this reduction for notational convenience only, a more general discussion will require the full notation. This labeling is graphically represented as

where the wiggly lines represent the intramolecular bond vectors, the squares and circles represent different atomic species, and the dashed lines represent the intermolecular site-site interactions of Eq. (3). The vector expressions of Eqs. (3) and (4) are easily written in the angular coordinates of the two molecules as

$$\begin{aligned}
U(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) &= u_{\alpha\gamma}(\mathbf{r}_{\alpha\gamma}) + u_{\gamma\gamma}(\mathbf{r}_{\alpha\gamma}, \Omega_1) + u_{\alpha\alpha}(\mathbf{r}_{\alpha\gamma}, \Omega_2) \\
&\quad + u_{\gamma\alpha}(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) \\
&= u^o(\mathbf{r}_{\alpha\gamma}) + u^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + u^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) \\
&\quad + u^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2),
\end{aligned} \tag{5}$$

where in the second line we make the natural notational correspondence of the o , l , r , b labeling of Chandler, Silbey, and Ladanyi¹² (CSL) to here indicate that o or none means that there is no angular dependence, l or left means that the term is dependent on the orientation of the left, or “1,” molecule, r or right means that the term is dependent on the orientation of the right, or “2,” molecule, and b or both means that the term is dependent on the orientation of both molecules.

Given this separation of the potential terms, we expand the distribution functions. The closure becomes

$$\begin{aligned}
c^o(\mathbf{r}_{\alpha\gamma}) &= e^{-\beta u^o(\mathbf{r}_{\alpha\gamma}) + \rho^o(\mathbf{r}_{\alpha\gamma}) + d^o(\mathbf{r}_{\alpha\gamma})} - t^o(\mathbf{r}_{\alpha\gamma}) - 1, \\
c^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) &= e^{-\beta u^o(\mathbf{r}_{\alpha\gamma}) + \rho^o(\mathbf{r}_{\alpha\gamma}) + d^o(\mathbf{r}_{\alpha\gamma})} (e^{-\beta u^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + d^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)} - 1) - t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1), \\
c^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) &= e^{-\beta u^o(\mathbf{r}_{\alpha\gamma}) + \rho^o(\mathbf{r}_{\alpha\gamma}) + d^o(\mathbf{r}_{\alpha\gamma})} (e^{-\beta u^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + t^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + d^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)} - 1) - t^r(\mathbf{r}_{\alpha\gamma}, \Omega_2), \\
c^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) &= e^{-\beta u^o(\mathbf{r}_{\alpha\gamma}) + \rho^o(\mathbf{r}_{\alpha\gamma}) + d^o(\mathbf{r}_{\alpha\gamma})} (\exp[-\beta u^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + t^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + d^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) - \beta u^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) \\
&\quad + t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + d^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) - \beta u^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + t^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + d^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)] - e^{-\beta u^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + d^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)} \\
&\quad - e^{-\beta u^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + t^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + d^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)} + 1) - t^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2).
\end{aligned} \tag{6}$$

Since this expansion is still expressed in terms of the full orientation-dependent molecular coordinate system, these are molecularly proper equations in the sense of the CSL theory. As we show below, as long as care is taken to properly account for certain subsets of diagrams through diagrammatic resummation or renormalization, a subset of exponential terms can be properly constructed for each radial site-site function.

The molecular OZ equation (1) expands in the various site-site origin labels as a site-site labeled block matrix, such that the submatrices $\mathbf{h}_{\alpha\gamma}$ and $\mathbf{c}_{\alpha\gamma}$ are defined as

$$\begin{aligned}
\mathbf{h}_{\alpha\gamma} &= \begin{pmatrix} h^o(k_{\alpha\gamma}) & h^r(k_{\alpha\gamma}, \Omega_2) \\ h^l(k_{\alpha\gamma}, \Omega_1) & h^b(k_{\alpha\gamma}, \Omega_1, \Omega_2) \end{pmatrix}, \\
\mathbf{c}_{\alpha\gamma} &= \begin{pmatrix} c^o(k_{\alpha\gamma}) & c^r(k_{\alpha\gamma}, \Omega_2) \\ c^l(k_{\alpha\gamma}, \Omega_1) & c^b(k_{\alpha\gamma}, \Omega_1, \Omega_2) \end{pmatrix},
\end{aligned} \tag{7}$$

and ρ is now block diagonal, with $\rho_{\alpha\gamma}=0$ in all terms when $\alpha \neq \gamma$ and, for this diatomic case,

$$\rho_{\alpha\alpha} = \begin{pmatrix} \rho/2 & \rho/2 \\ \rho/2 & \rho/2 \end{pmatrix}. \tag{8}$$

This expansion can be used to diagrammatically resum, or angularly renormalize, the intermolecular distributions in a manner similar to the methods used to treat long-ranged Coulomb potentials in simple fluids.^{16,23,24} For example, we consider the expansion of the left molecular diagrams. In the standard diagrammatic notation,^{16,25} the left indirect correlation function for a site-site fluid may contain

$$t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) = \begin{array}{c} \text{Diagram 1: A square node labeled } 1_\gamma \text{ at the top, connected by a wavy line to a circle node labeled } 1_\alpha \text{ at the bottom left, and by a straight line to a square node labeled } 2_\gamma \text{ at the bottom right.} \\ \text{Diagram 2: A square node labeled } 1_\gamma \text{ at the top, connected by a wavy line to a circle node labeled } 1_\alpha \text{ at the bottom left, and by a diagonal line to a square node labeled } 2_\gamma \text{ at the bottom right.} \end{array} + \dots \tag{9}$$

where the solid lines represent site-site Mayer f -bonds, the wiggly bonds represent the rigid, nondissociable intramolecular bonds, circles and squares denote different site species, filled circles and squares represent integrated site position variables, and open circles and squares represent unintegrated site positions. Notice that in this expansion we have not integrated over the 1_γ site or, equivalently, any of the configuration angles of molecule 1. We make the distinction between subsets of diagrams in $t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)$, which are dependent on only the site in the left molecule labeled 1_γ and those terms which depend on both sites 1_α and 1_γ . We will renormalize the closure by collecting the terms which are exclusively 1_γ dependent in $t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)$. We note that, by a simple coordinate change, these diagrams are equivalent to the none indirect correlation function between sites 1_γ and 2_γ , $t_{\gamma\gamma}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1|)$ as determined from the site origin mixture equations, with the $\gamma\gamma$ labeling here used in the same sense as above for the potential decomposition. Specifically, we change the coordinate system of $t^o(r_{\gamma\gamma}) \rightarrow t_{\gamma\gamma}^o(\mathbf{r}_{\alpha\gamma}, \Omega_1) = t_{\gamma\gamma}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1|)$ through the coordinate shift $\mathbf{r}_{\gamma\gamma} = \mathbf{r}_{\alpha\gamma} - \mathbf{l}_1$. Similarly, $t^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)$ can be renormalized by $t_{\alpha\alpha}^o(|\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2|)$, so that the resummed left and right closure equations, labeled $\bar{c}$, become

$$\begin{aligned} c^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1) &= G^o(r_{\alpha\gamma}) [e^{-u^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1) + \tau^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + d^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)} - 1] \\ &\quad - \tau^l(\mathbf{r}_{\alpha\gamma}, \Omega_1), \\ c^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2) &= G^o(r_{\alpha\gamma}) [e^{-u^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2) + \tau^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + d^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)} - 1] \\ &\quad - \tau^r(\mathbf{r}_{\alpha\gamma}, \Omega_2), \end{aligned} \quad (10)$$

which, together with $G^o(r_{\alpha\gamma}) = \exp(-\beta u^o(r_{\alpha\gamma}) + t^o(r_{\alpha\gamma}) + d^o(r_{\alpha\gamma}))$,

$u^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1) = u^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) - t_{\gamma\gamma}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1|)$, $u^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2) = u^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) - t_{\alpha\alpha}^o(|\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2|)$, and

$$\begin{aligned} \tau^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) &\equiv t^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) - t_{\gamma\gamma}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1|), \\ \tau^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) &\equiv t^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) - t_{\alpha\alpha}^o(|\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2|), \\ c^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) &= c^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1) - t_{\gamma\gamma}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1|) \\ &= c^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1) - \varphi^l(\mathbf{r}_{\alpha\gamma}, \Omega_1), \\ c^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) &= c^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2) - t_{\alpha\alpha}^o(|\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2|) \\ &= c^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2) - \varphi^r(\mathbf{r}_{\alpha\gamma}, \Omega_2), \end{aligned} \quad (11)$$

give the complete expansion for the c functions and define $\varphi^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)$ and $\varphi^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)$. These equations are analogous in form to the short-range form resulting from the renormalization of the Coulomb potential in simple fluid theory.²³ In the “both” case, we similarly use the none, left, and right terms to renormalize the both terms, such that

$$\begin{aligned} c^{-b}(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) &= G^o(r_{\alpha\gamma}) (\exp[-\beta u^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + t_{\gamma\alpha}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1 + \mathbf{l}_2|) + \tau_{\alpha\alpha}^l(\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2, \Omega_1') + \tau_{\gamma\gamma}^r(\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1, \Omega_2') + \tau^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) \\ &\quad + d^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)] - G^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) - G^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + 1) - \tau^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) \\ &= G^o(r_{\alpha\gamma}) [e^{-u^{-b}(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + \tau^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) + d^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)} G^l G^r - G^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) - G^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + 1] - \tau^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2), \end{aligned} \quad (12)$$

and

$$\begin{aligned} \tau^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) &\equiv t^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) - \varphi^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2), \\ c^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) &= c^{-b}(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) - \varphi^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2), \\ \varphi^b(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) &= t_{\gamma\alpha}^o(|\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1 + \mathbf{l}_2|) + \tau_{\alpha\alpha}^l(\mathbf{r}_{\alpha\gamma} + \mathbf{l}_2, \Omega_1') \\ &\quad + \tau_{\gamma\gamma}^r(\mathbf{r}_{\alpha\gamma} - \mathbf{l}_1, \Omega_2'), \\ G^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) &= e^{-u^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1) + \tau^l(\mathbf{r}_{\alpha\gamma}, \Omega_1) + d^l(\mathbf{r}_{\alpha\gamma}, \Omega_1)}, \\ G^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) &= e^{-u^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2) + \tau^r(\mathbf{r}_{\alpha\gamma}, \Omega_2) + d^r(\mathbf{r}_{\alpha\gamma}, \Omega_2)}. \end{aligned} \quad (13)$$

Here, we use Ω'_i as part of the generalized coordinate change, e.g.,

$$\tau^l(r_{\gamma\alpha}, \Omega_1) \rightarrow \tau^l_{\gamma\alpha}(r_{\gamma\alpha} + l, \Omega'_1), \quad (14)$$

which is necessary for the statement of the τ^l and τ^r functions in the both coordinate system. The explicit form of the coordinate change will in general be dependent on the particular choice of coordinate representation. The τ^l and τ^r functions, and not the t^l and t^r functions, are an appropriate choice for resummation in the definition of c^b so as to avoid overcounting many classes of diagrams.

The premise of this approach is succinctly stated in diagrammatic terms. Specifically, we have taken advantage of the equivalence

$$\equiv \hat{\mathbf{L}} \quad (15)$$

where by the operator $\hat{\mathbf{L}}$ we mean the operation of writing the diagram on the right in the coordinate system on the left by inverting the left bond, which is formally equivalent to changing the coordinate system to coincide with either the circle or square site, respectively. That is, for $f_1(r_{\alpha\gamma}, \Omega_1)$ the left diagram, $f_2(r_{\gamma\gamma})$ the radially symmetric right diagram before the inversion (equivalently, with the coordinate system coincident with but not rotationally fixed to the square), and $\mathbf{l}$ the intramolecular displacement vector between the circle and square sites, we have that

$$f_1(r_{\alpha\gamma}, \Omega_1) = f_2(|r_{\gamma\gamma} - \mathbf{l}|). \quad (16)$$

Further, the separation into o , l , r , b terms allows us to renormalize without violating the order of integration, i.e., it maintains the diagrammatically proper construction throughout. Below we show that this renormalization leads directly to the radial PISM equations, and by extension, to an entirely equivalent set of orientation-dependent equations generated from the PISM OZ equations.

B. Low order angular expansion of the site-site equations

Here, we show that the proper site-site integral equation theory is equivalent to a linearized, minimal basis set approximation to the full molecular theory. From the renormalized closure equations above, we linearize with respect to the τ^l , τ^r , and τ^b functions, and expand the l , r , and b e -bonds, i.e., terms such as $\exp(-\tilde{u}^l)$. Ignoring the cross terms and linearizing, we have

$$\begin{aligned} h^o(r_{\alpha\gamma}) &= e^{-\beta u(r_{\alpha\gamma}) + r^o(r_{\alpha\gamma}) + d^o(r_{\alpha\gamma})} - 1, \\ h^l(r_{\alpha\gamma}, \Omega_1) &= G^o(r_{\alpha\gamma}) [e^{-u^l(r_{\alpha\gamma}, \Omega_1)} - 1 + \tau^l(r_{\alpha\gamma}, \Omega_1) + d^l(r_{\alpha\gamma}, \Omega_1)], \\ h^r(r_{\alpha\gamma}, \Omega_2) &= G^o(r_{\alpha\gamma}) [e^{-u^r(r_{\alpha\gamma}, \Omega_2)} - 1 + \tau^r(r_{\alpha\gamma}, \Omega_2) + d^r(r_{\alpha\gamma}, \Omega_2)], \\ h^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) &= G^o(r_{\alpha\gamma}) [e^{-u^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) - u^l(r_{\alpha\gamma}, \Omega_1) - u^r(r_{\alpha\gamma}, \Omega_2)} \\ &\quad + \tau^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) + d^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) \\ &\quad + (\tau^l(r_{\alpha\gamma}, \Omega_1) + d^l(r_{\alpha\gamma}, \Omega_1))(\tau^r(r_{\alpha\gamma}, \Omega_2) \\ &\quad + d^r(r_{\alpha\gamma}, \Omega_2)) - e^{-u^l(r_{\alpha\gamma}, \Omega_1)} - e^{-u^r(r_{\alpha\gamma}, \Omega_2)} \\ &\quad + 1]. \end{aligned} \quad (17)$$

Now, in the terminology of Blum and Torruella¹ restricted to linear diatomic molecules, $h(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)$ may be expanded in rotational invariants Φ^{mnl} , such that

$$h(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) = \sum_{mnl} h^{mnl}(r_{\alpha\gamma}) \Phi^{mnl}(\Omega_1, \Omega_2, R_{\alpha\gamma}), \quad (18)$$

where the Φ^{mnl} functions are usually defined in terms of either of two generalized spherical harmonics conventions (Edmonds²⁶ or Messiah²⁷), m, n, l are the indices of the three rotation matrices or rotation angles, $R_{\alpha\gamma}$ is the unit vector along the displacement, $\mathbf{r}_{\alpha\gamma}$, between molecules in the $\alpha\gamma$ system, and the $h^{mnl}(r_{\alpha\gamma})$ are the coefficients of (projections on) the Φ^{mnl} expansion. If the molecular functions are expanded in this form, then the coefficients in the expansion of h or c may be determined via the molecular OZ equation. In the Edmonds phase and normalization convention¹

$$h^{mnl}(k) - c^{mnl}(k) = \rho \sum_{n_1 l_1 l_2} \frac{(2l+1)}{(2n_1+1)} (-1)^{m+n+n_1} \left\{ \begin{matrix} l_2 & l_1 & l \\ m & n & n_1 \end{matrix} \right\} \times \left(\begin{matrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{matrix} \right) h^{mn_1 l_1}(k) c^{n_1 n l_2}(k), \quad (19)$$

where the $h^{mnl}(k)$ and $c^{mnl}(k)$ functions are the Hankel transforms of $h^{mnl}(r)$ and $c^{mnl}(r)$, respectively, defined as

$$\begin{aligned} h^{mnl}(k) &= 4\pi i^l \int_0^\infty r^2 j_l(kr) h^{mnl}(r) dr, \\ c^{mnl}(k) &= 4\pi i^l \int_0^\infty r^2 j_l(kr) c^{mnl}(r) dr, \end{aligned} \quad (20)$$

with the $j_l(kr)$ the spherical Bessel functions, the term in parentheses, $\left(\begin{matrix} l_2 & l_1 & l \\ m & n & n_1 \end{matrix} \right)$, in Eq. (19) is the Wigner $3j$ -symbol,²⁶ and the term in curly braces, $\left\{ \begin{matrix} l_2 & l_1 & l \\ m & n & n_1 \end{matrix} \right\}$, is the $6j$ -symbol. If we now take the minimal, or isotropic, case, $h(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) \approx h^{000}(r_{\alpha\gamma})$, then Eq. (19) transformed and expanded in the site coordinate systems reduces to the matrix equation in Fourier space

$$\begin{pmatrix} h_{\alpha\alpha} & h_{\alpha\gamma} \\ h_{\gamma\alpha} & h_{\gamma\gamma} \end{pmatrix} = \begin{pmatrix} c_{\alpha\alpha} & c_{\alpha\gamma} \\ c_{\gamma\alpha} & c_{\gamma\gamma} \end{pmatrix} + \begin{pmatrix} c_{\alpha\alpha} & c_{\alpha\gamma} \\ c_{\gamma\alpha} & c_{\gamma\gamma} \end{pmatrix} \begin{pmatrix} \rho_{\alpha\alpha} & 0 \\ 0 & \rho_{\gamma\gamma} \end{pmatrix} \times \begin{pmatrix} h_{\alpha\alpha} & h_{\alpha\gamma} \\ h_{\gamma\alpha} & h_{\gamma\gamma} \end{pmatrix}, \quad (21)$$

where the submatrices are defined as

$$\begin{aligned} h_{\alpha\gamma} &= \begin{pmatrix} h^{a,000}(k_{\alpha\gamma}) & h^{r,000}(k_{\alpha\gamma}) \\ h^{l,000}(k_{\alpha\gamma}) & h^{b,000}(k_{\alpha\gamma}) \end{pmatrix}, \\ c_{\alpha\gamma} &= \begin{pmatrix} c^{a,000}(k_{\alpha\gamma}) & c^{r,000}(k_{\alpha\gamma}) \\ c^{l,000}(k_{\alpha\gamma}) & c^{b,000}(k_{\alpha\gamma}) \end{pmatrix}, \end{aligned} \quad (22)$$

and $\rho_{\alpha\gamma} = 0$ if $\alpha \neq \gamma$,

$$\rho_{\alpha\alpha} = \begin{pmatrix} \rho^{-o} & \rho^{-r} \\ \rho^{-l} & \rho^{-b} \end{pmatrix} = \begin{pmatrix} \rho/2 & \rho/2 \\ \rho/2 & \rho/2 \end{pmatrix} \quad (23)$$

otherwise. The closure equations in this approximation can be defined by noting that, as in any basis set expansion, any particular coefficient $h^{mnl}(r_{\alpha\gamma}) \propto \int d\Omega_1 d\Omega_2 h(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) \Phi^{mnl}$, with the proportionality constant given as $\Omega^2 = \int d\Omega_1 d\Omega_2 (\Phi^{mnl})^2$. In the minimal basis, this has the convenient form

$$h^{000}(r_{\alpha\gamma}) = \langle h(r_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_1 \Omega_2} \\ = \Omega^{-2} \int d\Omega_1 d\Omega_2 h(r_{\alpha\gamma}, \Omega_1, \Omega_2), \quad (24)$$

which gives, upon o, l, r, b decomposition, the closure equations

$$\begin{aligned} h^o(r_{\alpha\gamma}) &= e^{-\beta u^o(r_{\alpha\gamma}) + r^o(r_{\alpha\gamma}) + d^o(r_{\alpha\gamma})} - 1, \\ h^{l,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) \langle e^{-u^l(r_{\alpha\gamma}, \Omega_1)} - 1 + \tau^{l,000}(r_{\alpha\gamma}) \\ &\quad + d^l(r_{\alpha\gamma}, \Omega_1) \rangle_{\Omega_1}, \\ h^{r,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) \langle e^{-u^r(r_{\alpha\gamma}, \Omega_2)} - 1 + \tau^{r,000}(r_{\alpha\gamma}) \\ &\quad + d^r(r_{\alpha\gamma}, \Omega_2) \rangle_{\Omega_2}, \\ h^{b,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) [\langle e^{-u^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) - u^l(r_{\alpha\gamma}, \Omega_1) - u^r(r_{\alpha\gamma}, \Omega_2)} \\ &\quad + \tau^{b,000}(r_{\alpha\gamma}) + d^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_1, \Omega_2} \\ &\quad - \langle e^{-u^l(r_{\alpha\gamma}, \Omega_1)} \rangle_{\Omega_1} - \langle e^{-u^r(r_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_2} \\ &\quad + (\tau^{l,000}(r_{\alpha\gamma}) + d^{l,000}(r_{\alpha\gamma}))(\tau^{r,000}(r_{\alpha\gamma}) \\ &\quad + d^{r,000}(r_{\alpha\gamma})) + 1], \end{aligned} \quad (25)$$

together with the renormalizations

$$\begin{aligned} c^{l,000}(r_{\alpha\gamma}) &= c^{-l,000}(r_{\alpha\gamma}) - \langle \varphi^l(r_{\alpha\gamma}, \Omega_1) \rangle_{\Omega_1} \\ &= h^{l,000}(r_{\alpha\gamma}) - \tau^{l,000}(r_{\alpha\gamma}) - \langle t_{\gamma\gamma}^o(|r_{\alpha\gamma} - l_1|) \rangle_{\Omega_1}, \\ c^{r,000}(r_{\alpha\gamma}) &= c^{-r,000}(r_{\alpha\gamma}) - \langle \varphi^r(r_{\alpha\gamma}, \Omega_2) \rangle_{\Omega_2} \\ &= h^{r,000}(r_{\alpha\gamma}) - \tau^{r,000}(r_{\alpha\gamma}) - \langle t_{\alpha\alpha}^o(|r_{\alpha\gamma} + l_2|) \rangle_{\Omega_2}, \\ c^{b,000}(r_{\alpha\gamma}) &= c^{-b,000}(r_{\alpha\gamma}) - \langle \varphi^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_1, \Omega_2} \\ &= h^{b,000}(r_{\alpha\gamma}) - \tau^{b,000}(r_{\alpha\gamma}) - \langle t_{\gamma\alpha}^o(|r_{\alpha\gamma} - l_1 + l_2|) \\ &\quad + \tau_{\alpha\alpha}^{l,000}(|r_{\alpha\gamma} + l_2|) + \tau_{\gamma\gamma}^{r,000}(|r_{\alpha\gamma} - l_1|) \rangle_{\Omega_1, \Omega_2}. \end{aligned} \quad (26)$$

The forms of ϕ may be checked against the PISM equations by noting that, by definition, if $S_{\alpha\gamma}(r) = \delta(|r-l|)/(4\pi l^2 \rho)$ is the intramolecular distribution function as normally defined for rigid molecules, then,

$$\begin{aligned} x_{\alpha\gamma}(r) &= S_{\alpha\gamma} * \rho c_{\gamma\gamma}^o \\ &= \langle e^{-\beta u_{\gamma\gamma}(|r_{\alpha\gamma} - l_1|) + t_{\gamma\gamma}^o(|r_{\alpha\gamma} - l_1|)} \\ &\quad - t_{\gamma\gamma}^o(|r_{\alpha\gamma} - l_1|) - 1 \rangle_{\Omega_1} \\ &= \langle e^{-\beta u^l(r_{\alpha\gamma}, \Omega_1)} - \varphi^l(r_{\alpha\gamma}, \Omega_1) - 1 \rangle_{\Omega_1}, \end{aligned} \quad (27)$$

where the star denotes the usual convolution product. This, together with similar identities for the other functions, leads to the form derived by Rossky and Chiles for the proper CSL (Ref. 12) OZ equations.

To fully recover the h 's in Eq. (25) as the ‘‘HNC-like’’ results of Rossky and Chiles¹³ requires two things. First, we must reform ρ in Eq. (21) such that

$$\rho_{\alpha\alpha} = \begin{pmatrix} \rho & \rho \\ \rho & 0 \end{pmatrix}. \quad (28)$$

This deserves some comment.

First, the elimination of the factor 1/2 in ρ is not *ad hoc*. In particular, the double counting associated with the split of the original molecular functions of Eq. (1) into site-site coordinates

occurs through the ρ^b element alone. Elimination of this element eliminates the doubly counted diagrams which required the 1/2 factor.

To see this, consider the diagram pair,

(29)

which is generated from two different source terms in Eq. (1), i.e., the diagram on the left is generated by the $c^o(r_{ij}) * \rho_{jj}^o h^o(r_{jj})$ term, while the diagram on the right, which is obviously equivalent due to the normalization of the molecular S -bond, is generated from the $c^r(r_{ij}, \Omega_2) * \rho_{jj}^b h^l(r_{jj}, \Omega_1)$ source term in the OZ equations. Note that this particular combination occurs because the diagrams

(30)

are elements of $c(r, \Omega_1, \Omega_2)$ in the molecular expansion, whereas in the site-site systems, these diagrams are generated in the site-site OZ equations, and necessarily elements of t . Our renormalization assures that these diagrams occur only in φ , not $\mathfrak{c}$ or τ .

Similar diagram pairs are generated simultaneously in all of the elements of h when the ρ^b elements are included in the OZ matrix, which shows the necessity of the $\frac{1}{2}$ factor in the full ρ matrix. However, by eliminating the ρ^b element, each source term in h becomes unique, and thus the $\frac{1}{2}$ factor is eliminated. This factor will reoccur below, however, when we reintroduce a subset of the source terms contributed via the ρ^b elements; in that case, the $\frac{1}{2}$ factor remains necessary due to doubly counting a certain class of site-site bridge diagrams.

Second, elimination of the ρ^b element together with the linearization of the closure forms a closed set, in that the diagrams which are included may be calculated exactly, i.e., there is no basis set truncation error associated with the diagrams of the minimal angular basis set for those diagrams explicitly included.

Finally, for the HNC-like approximation, the molecular bridge functions must be set to 0, $d^o=d^l=d^r=d^b=0$, and we must subtract one additional site-site function from $h^{b,000}$ [Eq. (25)], such that

$$\begin{aligned}
 h_{\text{HNC}}^{b,000}(r_{\alpha\gamma}) = & G^o(r_{\alpha\gamma}) [\langle e^{-u^b(r_{\alpha\gamma}, \Omega_1, \Omega_2)} - u^{-l}(r_{\alpha\gamma}, \Omega_1) - u^{-r}(r_{\alpha\gamma}, \Omega_2) \\
 & + \tau^{b,000}(r_{\alpha\gamma}) + d^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_1, \Omega_2} \\
 & - \langle (e^{-u^b(r_{\alpha\gamma}, \Omega_1, \Omega_2)} - 1)(e^{-u^{-l}(r_{\alpha\gamma}, \Omega_1)} - 1) \\
 & \times (e^{-u^{-r}(r_{\alpha\gamma}, \Omega_2)} - 1) \rangle_{\Omega_1, \Omega_2} - \langle e^{-u^{-l}(r_{\alpha\gamma}, \Omega_1)} \rangle_{\Omega_1} \\
 & - \langle e^{-u^{-r}(r_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_2} + (\tau^{l,000}(r_{\alpha\gamma}) + d^{l,000}(r_{\alpha\gamma})) \\
 & \times (\tau^{r,000}(r_{\alpha\gamma}) + d^{r,000}(r_{\alpha\gamma}) + 1)].
 \end{aligned}
 \tag{31}$$

For molecules with more than two sites, there would necessarily be similar terms subtracted from $h^{l,000}$ and $h^{r,000}$, but for diatomics this extension is not necessary.

C. An exact site-renormalized theory for the PISM-OZ equations

The linearization used in generating the diagrammatically proper interaction site formalism directly from the molecular theory is not strictly necessary. Consider Eq. (11). If we truncate at the isotropic basis, we have

$$\begin{aligned} h^{l,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) (\langle e^{-u^{-l}(r_{\alpha\gamma}, \Omega_1) + d^l(r_{\alpha\gamma}, \Omega_1) + \tau^{l,000}(r_{\alpha\gamma})} \rangle_{\Omega_1} - 1) \\ &= G^o(r_{\alpha\gamma}) (\langle e^{-u^{-l}(r_{\alpha\gamma}, \Omega_1) + d^l(r_{\alpha\gamma}, \Omega_1)} \rangle_{\Omega_1} e^{\tau^{l,000}(r_{\alpha\gamma})} - 1). \end{aligned} \quad (32)$$

That is, $\exp(\tau^{l,000}(r_{\alpha\gamma}))$ is independent of the angular integrations, by construction, and thus can be factored exactly from the closure statements. We can take advantage of this in the proper formulation by suitably renormalizing the proper OZ site-site matrix equations. In the notation of Rossky and Chiles¹³ together with the modified, screened density approach recently proposed¹⁷, the proper site-site matrix equations in Fourier space are

$$\mathbf{H} = \mathbf{C} + \mathbf{C} \rho \mathbf{H}, \quad (33)$$

where

$$\begin{aligned} \rho_{\alpha\alpha} &= \begin{pmatrix} \rho & \eta \\ \eta & 0 \end{pmatrix}, \\ \mathbf{H}_{\alpha\gamma} &= \begin{pmatrix} h^o(k_{\alpha\gamma}) & h^{r,000}(k_{\alpha\gamma}) \\ h^{l,000}(k_{\alpha\gamma}) & h^{b,000}(k_{\alpha\gamma}) + S(k_{\alpha\gamma}) \end{pmatrix}, \\ \mathbf{C}_{\alpha\gamma} &= \begin{pmatrix} c^o(k_{\alpha\gamma}) & c^{r,000}(k_{\alpha\gamma}) \\ c^{l,000}(k_{\alpha\gamma}) & c^{b,000}(k_{\alpha\gamma}) + S(k_{\alpha\gamma}) \end{pmatrix}, \\ S(r_{\alpha\gamma}) &= \frac{\delta(k_{\alpha\gamma} - l_{\alpha\gamma})}{4\pi l_{\alpha\gamma}^2 \eta}, \end{aligned} \quad (34)$$

$l_{\alpha\gamma}$ is the intramolecular bond length, $S(r)$ is the (rigid) intramolecular distribution, and we use η in the general sense, whether as the normalization appropriate to reactive fluids^{28–30} or as a bridge function approximation variable.¹⁷ If we now renormalize such that

$$\begin{aligned} \tau_{\alpha\gamma} &= \mathbf{H}_{\alpha\gamma} - \mathbf{C}_{\alpha\gamma} - \varphi_{\alpha\gamma} = \mathbf{H}_{\alpha\gamma} - \mathbf{C}_{\alpha\gamma} \\ &= \begin{pmatrix} t^o(k_{\alpha\gamma}) & \tau^{r,000}(k_{\alpha\gamma}) \\ \tau^{l,000}(k_{\alpha\gamma}) & \tau^{b,000}(k_{\alpha\gamma}) \end{pmatrix}, \end{aligned} \quad (35)$$

that is, we restrict the functions from the site-site equations to the τ set from Eqs. (11)–(13), then we close the system by

$$\begin{aligned} c^{-o}(r_{\alpha\gamma}) &= c^o(r_{\alpha\gamma}) = e^{-\beta u(r_{\alpha\gamma}) + \rho^o(r_{\alpha\gamma}) + d^o(r_{\alpha\gamma})} - t^o(r_{\alpha\gamma}) - 1, \\ c^{-l,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) (\langle e^{-u^{-l}(r_{\alpha\gamma}, \Omega_1) + d^l(r_{\alpha\gamma}, \Omega_1)} \rangle_{\Omega_1} e^{\tau^{l,000}(r_{\alpha\gamma})} - 1) - \tau^{l,000}(r_{\alpha\gamma}), \\ c^{-r,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) (\langle e^{-u^{-r}(r_{\alpha\gamma}, \Omega_2) + d^r(r_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_2} e^{\tau^{r,000}(r_{\alpha\gamma})} - 1) - \tau^{r,000}(r_{\alpha\gamma}), \\ c^{-b,000}(r_{\alpha\gamma}) &= G^o(r_{\alpha\gamma}) (\langle e^{-u^{-b}(r_{\alpha\gamma}, \Omega_1, \Omega_2) - u^{-l}(r_{\alpha\gamma}, \Omega_1) - u^{-r}(r_{\alpha\gamma}, \Omega_2) + d^b(r_{\alpha\gamma}, \Omega_1, \Omega_2) + d^l(r_{\alpha\gamma}, \Omega_1) + d^r(r_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_1, \Omega_2} e^{\tau^{b,000}(r_{\alpha\gamma}) + \tau^{l,000}(r_{\alpha\gamma}) + \tau^{r,000}(r_{\alpha\gamma})} \\ &\quad - \langle e^{-u^{-l}(r_{\alpha\gamma}, \Omega_1) + d^l(r_{\alpha\gamma}, \Omega_1)} \rangle_{\Omega_1} e^{\tau^{l,000}(r_{\alpha\gamma})} - \langle e^{-u^{-r}(r_{\alpha\gamma}, \Omega_2) + d^r(r_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_2} e^{\tau^{r,000}(r_{\alpha\gamma})} + 1) - \tau^{b,000}(r_{\alpha\gamma}), \end{aligned} \quad (36)$$

and the renormalization functions are

$$\begin{aligned}
\varphi^{l,000}(r_{\alpha\gamma}) &= \langle e^{-u^{-l}(r_{\alpha\gamma},\Omega_1)+d^l(r,\Omega_1)} \rangle_{\Omega_1} - 1, \\
\varphi^{r,000}(r_{\alpha\gamma}) &= \langle e^{-u^{-r}(r_{\alpha\gamma},\Omega_2)+d^r(r_{\alpha\gamma},\Omega_2)} \rangle_{\Omega_2} - 1, \\
\varphi^{b,000}(r_{\alpha\gamma}) &= \langle e^{-u^{-b}(r_{\alpha\gamma},\Omega_1,\Omega_2)+d^b(r_{\alpha\gamma},\Omega_1,\Omega_2)-u^{-l}(r_{\alpha\gamma},\Omega_1)+d^l(r_{\alpha\gamma},\Omega_1)-u^{-r}(r_{\alpha\gamma},\Omega_2)+d^r(r_{\alpha\gamma},\Omega_2)} \rangle_{\Omega_1,\Omega_2} \\
&\quad - \langle (e^{-u^{-b}(r_{\alpha\gamma},\Omega_1,\Omega_2)+d^b(r_{\alpha\gamma},\Omega_1,\Omega_2)} - 1) \\
&\quad \times (e^{-u^{-l}(r_{\alpha\gamma},\Omega_1)+d^l(r_{\alpha\gamma},\Omega_1)} - 1)(e^{-u^{-r}(r_{\alpha\gamma},\Omega_2)+d^r(r_{\alpha\gamma},\Omega_2)} - 1) \rangle_{\Omega_1,\Omega_2} \\
&\quad - \langle (e^{-u^{-l}(r_{\alpha\gamma},\Omega_1)+d^l(r_{\alpha\gamma},\Omega_1)-u^{-r}(r_{\alpha\gamma},\Omega_2)+d^r(r_{\alpha\gamma},\Omega_2)} \rangle_{\Omega_1,\Omega_2}.
\end{aligned} \tag{37}$$

Though Eqs. (36) and (37) are labeled for the 000 coefficients, the form of these equations does not change for any particular $c^{mnl}(r)$ coefficient, since the only change then would be to sum over the various terms in τ^{mnl} in the exponentials rather than truncating at 000, and then take into account the Φ^{mnl} weighting in the angular average for each coefficient. Therefore, with this labeling caveat, Eqs. (36) and (37) are general and exact for any arbitrary mnl coefficient required. We also note that, by construction, we have, for the HNC approximation truncated at the isotropic basis,

$$\begin{aligned}
g(r_{\alpha\gamma},\Omega_1,\Omega_2) &= \exp(-\beta u(r_{\alpha\gamma}) + t^o(r_{\alpha\gamma}) - u^{-l}(r_{\alpha\gamma},\Omega_1) \\
&\quad + \tau^{l,000}(r_{\alpha\gamma}) - u^{-r}(r_{\alpha\gamma},\Omega_2) + \tau^{r,000}(r_{\alpha\gamma}) \\
&\quad - u^{-b}(r_{\alpha\gamma},\Omega_1,\Omega_2) + \tau^b(r_{\alpha\gamma})),
\end{aligned} \tag{38}$$

that is, we can generate $g(r_{\alpha\gamma},\Omega_1,\Omega_2)$ self-consistently from the site-site distribution functions through Eqs. (36) and (37). We will use this result below to calculate the diatomic analog of the 3D site-site distribution functions, $\langle g(r_{\alpha\gamma},\Omega_1,\Omega_2) \rangle_{\Omega_2}$.

We note that the last two diagrams on h -bonds in the definition of $\varphi^{b,000}$ [the second and third functions of the $\varphi^{b,000}$ definition in Eq. (37)] must be accounted for explicitly, since in the site-site theory the first is a bridge function and necessarily contained in $c^{-b,000}$ due to the construction of the OZ matrix, while the $l \times r$ multiplication for the second is a c^b diagram in the site-site construction. Further, we only need to restrict the HNC approximation for this closure to the molecular HNC, i.e., $d^o = d^l = d^r = d^b = 0$. Thus, unlike the HNC-like closure of PISM or XRISM, this closure is exact to order ρ^0 , since the intermolecular site-site bridge diagrams of order ρ^0 are included by construction. Generalization to arbitrary molecules preserves this exactness, since the generalization would be in additional (angular) dimensions, not diagrammatic terms.

Finally, we return the complete isotropic approximation to Eqs. (11)–(13) and (21), and make the system formally exact to order ρ^1 , by writing the proper site-site OZ equation as

$$\begin{aligned}
\mathbf{H} &= \chi + \mathbf{S} + (\chi + \mathbf{S})\rho\mathbf{H}, \\
\tau &= [\mathbf{I} - (\chi + \mathbf{S})\rho]^{-1}(\chi + \mathbf{S}) - \bar{\mathbf{c}} - \mathbf{S},
\end{aligned} \tag{39}$$

where χ is the proper chain sum²⁰ defined by

$$\chi = [\mathbf{I} - \mathbf{c}\rho']^{-1}\mathbf{c}, \tag{40}$$

ρ is the same as Eq. (34), $\mathbf{C} = \mathbf{c} + \mathbf{S} = \bar{\mathbf{c}} - \varphi + \mathbf{S}$, the components of $\bar{\mathbf{c}}$ are the Fourier transforms of Eq. (36), the φ are the Fourier transforms of the functions defined in Eq. (37) (and $\varphi^{o,000}=0$), and ρ' is defined in general as $\rho'_{ij}=0$ when $i \neq j$ and

$$\rho'_{ii} = \begin{pmatrix} 0 & 0 \\ 0 & \eta/2 \end{pmatrix} \tag{41}$$

otherwise.

That χ is diagrammatically proper, and that the $\frac{1}{2}$ factor in ρ' is necessary, is easily demonstrated. Consider first the convolution of two angular-dependent molecular f -bonds in the same conventions as Eq. (19),

$$\begin{aligned} \langle f(k_{13}, \Omega_1, \Omega_3) \rho f(k_{23}, \Omega_2, \Omega_3) \rangle_{\Omega_3} &= \sum_{n_1 l_1 l_2} \frac{2l+1}{2n_1+1} \\ &\times (-)^{m+n+n_1} \begin{Bmatrix} l_2 & l_1 & l \\ m & n & n_1 \end{Bmatrix} \begin{pmatrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{pmatrix} \\ &\times f^{mn l_1}(k) \rho f^{n_1 n l_2}(k) \Phi^{mnl}. \end{aligned} \quad (42)$$

If we now expand the molecular f -bonds in site-site potentials, then the site-site diagram

(43)

is clearly a site-site diagrammatic element of the set included in Eq. (42). This is a molecular convolution integral for rigid molecules, but if we use only site-site pair functions, this is a site-site bridge diagram. Now, consider the series in Eq. (42) truncated at $(m_1, n_1, l_1, l_2) = (0, 0, 0, 0)$,

$$\begin{aligned} \langle f(k_{13}, \Omega_1, \Omega_3) \rho f(k_{23}, \Omega_2, \Omega_3) \rangle_{\Omega_3} \\ \approx (2l+1)(-)^{m+n} \begin{Bmatrix} 0 & 0 & l \\ m & n & 0 \end{Bmatrix} \begin{pmatrix} 0 & 0 & l \\ 0 & 0 & 0 \end{pmatrix} \\ \times f^{m00}(k) \rho f^{0n0}(k) \Phi^{mnl} = f^{000}(k) \rho f^{000}(k). \end{aligned} \quad (44)$$

That is, the convolution product diagram

(45)

is then necessarily a member of both Eq. (42) and the truncated series in Eq. (44), and is the first term in the molecular convolution series expansion of the bridge diagram in Eq. (43). Again, as with the exponentiation in the closure, some convolution diagrams which are apparently improper are, in fact, required and a simple consequence of the topological expansion when the generalization to complete angular dependence is accounted for in a full molecular expansion.

Finally, the $\frac{1}{2}$ factor is reintroduced for ρ' only, due to the symmetry of the S -bond between field points in Eq. (43). Since the site-renormalized series expansions of $c^l(r_{ij}, \Omega_1)$, $c^l(r_{ij}, \Omega_2)$, and $c^b(r_{ij}, \Omega_1, \Omega_2)$ from the angular generalizations of Eqs. (36) and (37) collectively begin with the diagram series

(46)

respectively, the class of diagrams which can only be introduced by chaining **S**-bonds is generated only through Eq. (39), ensuring that each source term in Eqs. (39) and (40) generates a unique, different set of diagrams. However, the χ source equation reintroduces the site-site bridge diagrams, such as Eq. (43), which have *S*-bonds attached across field points and are generated by the molecular convolution integral. If we were using only one unique coordinate system, then only one of, for instance, the two diagrams such as

(47)

would be generated from the χ equation. Thus, since the χ equation generates all of the possible site-site pair functions simultaneously, and thus generates n equivalent pairs of these bridge diagrams, the $1/n$ factor, $1/2$ here for diatomics, remains for the χ series alone.

In general, $\chi(\mathbf{r}, \Omega_1, \Omega_2)$ componentwise in the Φ^{mnl} expansions is given by

$$\chi^{mnl}(k) = c^{mnl}(k) + \sum_{n_1 l_1 l_2} \frac{2l+1}{2n_1+1} (-)^{m+n+n_1} \begin{Bmatrix} l_2 & l_1 & l \\ m & n & n_1 \end{Bmatrix} \times \begin{pmatrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{pmatrix} c^{mn_1 l_1} \rho' \chi^{n_1 n l_2}.$$

(48)

Equation (48) is generated from **c**, not **C**, to avoid indiscriminately chaining **S**-bonds together, which would generate truly improper topological inconsistencies, as in the RISM theory, while still including those (fully connected site-site) molecular diagrams not generated directly by Eq. (39). In the site-site theory, then, χ is another necessary bridge diagram series, and in general, using the expansions of Eq. (48), the proper site-site renormalized OZ equation (39), and the complete Φ^{mnl} expansions for Eqs. (36) and (37), forms a proper site-site orientation-dependent molecular fluid integral equation theory formally equivalent to the molecular OZ theory of Blum and Torruella. It is thus formally exact, yet embeds the form of the PISM radial site-site theory.

This makes clear the difference in consequences of a derivable mathematical cluster theory for molecular fluids versus augmenting theories with physical cluster rules.¹² Notice above that those diagrams which involve the indiscriminant chaining of **S** with itself are still disallowed. Yet, there remain classes of diagrams which are consistent with the exponentiation and convolution of the rotational invariants that though they would at first appear to be topologically inconsistent (e.g., RISM) are allowed and required once we make the complete

generalization to full molecular rotational dimensionality. These real-space and convolution products contribute to the isotropic correlations even though each diagrammatic term may individually be anisotropic.

III. Results

Here, we present numerical results in the first, isotropic approximation to the theory outlined above. We have calculated the radial distribution functions in the molecular HNC approximation according to Eqs. (36), (37), (39), and (40) for three homonuclear Lennard-Jones diatomic models. For two of the models, $l^*=l/\sigma=0.729, 0.547$, we have used $\eta=\rho$. In the smallest bond length considered here, $l^*=0.329$, which case is a notoriously difficult test of theory, we were unable to obtain a stable solution with $\eta=\rho$, and thus took the approximation $\eta \approx \rho/(1+(\rho\pi d^4/3l))=\rho/2.668$ from our effective density treatment of homonuclear diatomics.

¹⁷ However, following the principle that a solution with η as close to ρ as possible is the best test of the theory as presented, by successive iteration, we were able to effectively minimize the difference between η and ρ and calculate a solution with $\eta=\rho/1.24$. We report both results below. Also, for an apolar Lennard-Jones model of HCl we used again $\eta=\rho$. Since a self-consistent construction for $g(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2)$ is integral to these calculations, we have also calculated the diatomic analog of the 3D site-site distribution functions, $\langle g(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_2}$. In the first section, we detail some of the numerical considerations necessary for these calculations. In the second section, we show the results for the homonuclear models, including comparisons of the site-site averages to our effective density variant of PISM and as measured from simulation. In the third section, we present results for the apolar HCl model in comparison to RISM, the standard construction of PISM, and simulation.

Since we published the effective density theory, subsequent analysis has found that the solution for η used in that and prior work³¹ was an asymptotic equation. In those cases, we used a hard sphere result $\eta \approx \rho/(1+\rho\pi d^4/3l)$. Further work shows that this result is a special case of the general solution, valid for $l \ll d$. However while this result is a special formal case, in the vast majority of bond lengths, it is still remarkably close to the general result which can be obtained by numerical integration.³² This is further demonstrated in the work using the effective density theory in extension to other closure approximations (work in progress).

A. Numerical methods

Though the angular-dependent terms in Eqs. (36) and (37) are topologically similar to the four-point bridge diagrams previously calculated,^{18,19,33} the rigid molecule approximation greatly simplifies these terms, so that the angular averages can be readily calculated using direct integration methods. We use the relative coordinate definitions $(\mathbf{r}_{\alpha\gamma}, \Omega_1, \Omega_2) = (r_{\alpha\gamma}, \theta_1, \theta_2, \phi_{12})$ throughout. We choose the intermolecular axis, z , aligned along $r_{\alpha\gamma}$, θ_1 and θ_2 the angles of the intramolecular bond vectors $\mathbf{l}_1$ and $\mathbf{l}_2$ with respect to z , and $\phi_{12}=\phi_1-\phi_2$ the dihedral angle between the two molecules. For the left and right terms we require the integrals

$$\begin{aligned} G^{l,000}(r_{\alpha\gamma}) &= \langle e^{-u^{-l}(\mathbf{r}_{\alpha\gamma}, \Omega_1)} \rangle_{\Omega_1} = \frac{1}{2} \int_0^\pi \sin\theta_1 d\theta_1 \\ &\quad \times e^{-\beta u_{\gamma\gamma}(\sqrt{r_{\alpha\gamma}^2 + l^2 - 2r_{\alpha\gamma}l\cos\theta_1}) + l_{\gamma\gamma}^o(\sqrt{r_{\alpha\gamma}^2 + l^2 - 2r_{\alpha\gamma}l\cos\theta_1})}, \\ G^{r,000}(r_{\alpha\gamma}) &= \langle e^{-u^{-r}(\mathbf{r}_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_2} = \frac{1}{2} \int_0^\pi \sin\theta_2 d\theta_2 \\ &\quad \times e^{-\beta u_{\alpha\alpha}(\sqrt{r_{\alpha\gamma}^2 + l^2 + 2r_{\alpha\gamma}l\cos\theta_2}) + l_{\alpha\alpha}^o(\sqrt{r_{\alpha\gamma}^2 + l^2 + 2r_{\alpha\gamma}l\cos\theta_2})}, \end{aligned} \quad (49)$$

where l is the bond length, $u_{\gamma\gamma}(|\mathbf{r}_{\alpha\gamma}-\mathbf{l}|)$ is the potential between sites γ on molecule 1 and γ on molecule 2 written in the $\alpha\gamma$ coordinate system, and similarly for the $l_{\gamma\gamma}^o$ terms. These integrals were calculated using a trapezoid rule, with 16 points in θ found to be sufficient for numerical convergence in the cases investigated here. For the both terms, we require

$$\begin{aligned}
G^{b,000}(r_{\alpha\gamma}) &= \langle e^{-u^{-b}(r_{\alpha\gamma}, \Omega_1, \Omega_2) - u^{-l}(r_{\alpha\gamma}, \Omega_2) - u^{-r}(r_{\alpha\gamma}, \Omega_2)} \rangle_{\Omega_1, \Omega_2} \\
&= \frac{1}{8\pi} \int_0^\pi \int_0^\pi \int_0^{2\pi} \sin \theta_1 d\theta_1 \sin \theta_2 d\theta_2 d\varphi_{12} \exp(-\beta u_{\gamma\alpha}(\sqrt{r_{\alpha\gamma}^2 + 2l^2 + 2r_{\alpha\gamma}l \cos \theta_2 - 2r_{\alpha\gamma}l \cos \theta_1 - 2l^2 \cos \gamma_{12}}) \\
&\quad + t_{\gamma\alpha}^0(\sqrt{r_{\alpha\gamma}^2 + 2l^2 + 2r_{\alpha\gamma}l \cos \theta_2 - 2r_{\alpha\gamma}l \cos \theta_1 - 2l^2 \cos \gamma_{12}}) + t_{\alpha\alpha}^{l,000}(\sqrt{r_{\alpha\gamma}^2 + l^2 + 2r_{\alpha\gamma}l \cos \theta_2}) \\
&\quad + t_{\gamma\gamma}^{r,000}(\sqrt{r_{\alpha\gamma}^2 + l^2 - 2r_{\alpha\gamma}l \cos \theta_1}) - \beta u_{\gamma\gamma}(\sqrt{r_{\alpha\gamma}^2 + l^2 - 2r_{\alpha\gamma}l \cos \theta_1}) + t_{\gamma\gamma}^0(\sqrt{r_{\alpha\gamma}^2 + l^2 - 2r_{\alpha\gamma}l \cos \theta_1}) \\
&\quad - \beta u_{\alpha\alpha}(\sqrt{r_{\alpha\gamma}^2 + l^2 + 2r_{\alpha\gamma}l \cos \theta_2}) + t_{\alpha\alpha}^0(\sqrt{r_{\alpha\gamma}^2 + l^2 + 2r_{\alpha\gamma}l \cos \theta_2})),
\end{aligned} \tag{50}$$

with $\cos \gamma_{12} \equiv \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos \varphi_{12}$, and the other terms required in c^b and the ϕ^b definitions have similar expansions. For the models examined here, we found sufficient convergence with 16 points in θ and 32 points in ϕ . The two-dimensional (2D) functions,

$$\langle g(r_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_2} = g(r_{\alpha\gamma}, \theta_1) = \frac{1}{4\pi} \int d\theta_2 d\varphi_{12} g(r_{\alpha\gamma}, \theta_1, \theta_2, \varphi_{12}), \tag{51}$$

where $g(r_{\alpha\gamma}, \theta_1, \theta_2, \varphi_{12})$ is taken from Eq. (38) in relative coordinates, are a straightforward extension of these integration types, and require only accumulation at the appropriate integration step in the full 3D integrals for the both functions. We also calculate the excess internal energy directly in the usual molecular method from $g(r_{\alpha\gamma}, \theta_1, \theta_2, \varphi_{12})$,

$$\begin{aligned}
\frac{\beta U^{\text{ex}}}{N} &= 2\pi\rho \int r_{\alpha\gamma}^2 dr_{\alpha\gamma} \langle \beta U(r_{\alpha\gamma}, \Omega_1, \Omega_2) g(r_{\alpha\gamma}, \Omega_1, \Omega_2) \rangle_{\Omega_1, \Omega_2} \\
&= \frac{1}{4}\rho \int r_{\alpha\gamma}^2 dr_{\alpha\gamma} \sin \theta_1 d\theta_1 \sin \theta_2 d\theta_2 d\varphi_{12} (\beta u_{\alpha\gamma}(r_{\alpha\gamma}) + \beta u_{\gamma\gamma}(\sqrt{r_{\alpha\gamma}^2 + l^2 - 2r_{\alpha\gamma}l \cos \theta_1}) + \beta u_{\alpha\alpha}(\sqrt{r_{\alpha\gamma}^2 + l^2 + 2r_{\alpha\gamma}l \cos \theta_2}) \\
&\quad + \beta u_{\gamma\alpha}(\sqrt{r_{\alpha\gamma}^2 + 2l^2 + 2r_{\alpha\gamma}l \cos \theta_2 - 2r_{\alpha\gamma}l \cos \theta_1 - 2l^2 \cos \gamma_{12}})) g(r_{\alpha\gamma}, \theta_1, \theta_2, \varphi_{12}),
\end{aligned} \tag{52}$$

rather than through the regular site-site expression. For the homonuclear cases, we previously reported $\beta U^{\text{ex}}/N\epsilon^*$ termwise¹⁷ and continue that practice here, which requires dividing the total energy result by $4\epsilon^*$, where $\epsilon^* = T^*-1$. For the heteronuclear case, we report the full value, $\beta U^{\text{ex}}/N$. In principle, though we do not carry out that analysis here, the pressure could also be calculated in closed form from the full molecular definitions, which has not yet proven satisfactory in the standard site-site methods.

In terms of iteration procedures, we treat this system in the same way as was done for atomic h -bonded bridge diagrams,^{18,19} in that the angular-averaged terms are held constant while iterating the closure and OZ equations to convergence, and then the new angular averages computed. In all cases considered here, machine-precision convergence for the equations required approximately 12 such total iterations of the angular-averaged functions. In these particular cases, we found no need to use relaxation techniques on the angular terms, though the standard Picard relaxation-iteration procedure for the OZ-closure cycle was necessary.

B. Homonuclear models

Here, we report the results of our calculations for three homonuclear diatomic systems. In all cases, we use the phase point defined in reduced units as $\rho^* = \rho\sigma^3 = 0.524$ and $T^* = k_B T / \epsilon = 2.20$, where σ and ϵ are the usual Lennard-Jones parameters. In the homonuclear cases, we report results using a radial grid of 512 points and a range of 15σ , except for both N_2 results, where a grid of 2048 points over a range of 60σ was employed. The three models are for “ N_2 ” or short bond length, $l^* = l/\sigma = 0.329$, “ Cl_2 ” or medium bond length, $l^* = 0.547$, and a long bond length model, $l^* = 0.729$. We use the approximation defined previously¹⁷ for η for the N_2 model only, with $\eta = \rho/2.668$ and $\eta = \rho/1.24$, which last was the effective density closest to $\eta = \rho$ for which we could calculate stable results through successive iteration. For the two longer bond lengths, we use $\eta = \rho$.

The $g^{000}(r)$ results are summarized in Fig. 1 for the homonuclear cases. In the first case, the structural results using $\eta=\rho/2.668$ are not significantly different than our previous calculations in the work on η for this difficult model geometry and phase point. However, the $\eta=\rho/1.24$ results are a significant qualitative and quantitative improvement. For the longer bond lengths, with $\eta=\rho$, the results are also an overall improvement over the previous results. These changes contribute to an improvement in the thermodynamic values over those previously reported for the η approximation. This is summarized in Tables I and II. Table I gives the results for the Kirkwood-Buff G integral, $G_{\alpha\gamma}=4\pi\int r^2 dr_{\alpha\gamma} h^{000}(r_{\alpha\gamma})$, and Table II gives the (termwise, ϵ^* reduced) excess internal energy. In these tables, η -HNC is the $\eta=\rho/2.668$ approximation with the HNC-like closure, while the results labeled mHNC are the results of this work using Eqs. (36), (37), (39), and (40), here with the $\eta=\rho/1.24$ approximation for N_2 and $\eta=\rho$ otherwise. The results for both thermodynamic quantities investigated are very encouraging, especially in comparison to the results found using previous theories.

The angular distributions for these models are summarized using two different perspectives in Figs. 2–4 for the short, medium, and long bond lengths, respectively. These are the 2D site-site distribution functions, $g(r_{\alpha\gamma}, \theta_1)$, calculated self-consistently according to Eqs. (38) and (51). Here, we note the fundamental asymmetries of the homonuclear diatomic system, with the first solvation shell being maximal close to $\theta_1=\pi/2$, or the x -like configuration, and gradually minimizing to $\theta_1=0$ and $\theta_2=\pi$, with $\theta_1=0$ being the most favored secondary orientation for all bond lengths. In these cases, we label the σ distance across the right hand figure in order to give perspective, and show that the solvation peaks at $\theta_1=0$ are pushed out by the bond distance, and thus the first shell is pushed out to $\sigma+l$ in the $\theta_1=0$ case for each bond length, while the $\theta_1=\pi$ first shell distance is approximately σ for each case. In the second shell, the effect of the bond length is also evident, with the N_2 model being the most asymmetric, and the region $\theta_1=\pi/2 \rightarrow \theta_1=0$ somewhat favored, while in the medium and long bond lengths the secondary solvation shell echoes the first solvation shell of the long bond length, with $\theta_1=\pi/2$ the most favored orientation, and the $\theta_1=\pi, 0$ orientations symmetrically the secondary orientations. Another interesting feature is the $\theta_1=\pi/4 \rightarrow 0$ region in the first solvation shell. Comparing the different bond lengths, it appears that a secondary orientation effect of the first solvation shell towards $\theta_1=0$ is present for all three bond lengths, so that an almost united-atom type “shared” solvation can occur between sites 1 and 2 on molecule 1 for the short and medium bond length models.

C. A heteronuclear model

Next, we consider the results for an apolar, heteronuclear model of HCl.³⁴ The presence of an essentially auxiliary site for the H atom makes this model especially interesting. The phase point we use here is $\rho=0.018$ molecules/ $\text{\AA}^3$ and $T=610$ K. We use the standard Lorentz-

Berthelot mixing rules, $\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}}$ and $\sigma_{ii} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj})$. The model parameters we use here are a bond length of $l=1.3$ $\text{\AA}$, $\sigma_{Cl}=3.353$ $\text{\AA}$, $\epsilon_{Cl}/k_B=259$ K, $\sigma_H=0.4$ $\text{\AA}$, and $\epsilon_H/k_B=20$ K. Strictly speaking, since we are using a complete potential description, the parameters for H are unnecessary in this apolar model. However, RISM and PISM both require an explicit potential for H to avoid collapse at the origin. So, as is a standard practice in integral equation theories, for both simulation and the integral equations, we include the H–H interactions. For emphasis and completeness, we also report results using the theory derived in this work for the HCl model with the H interaction set to 0. The thermodynamic and one-dimensional $g(r)$ results are indistinguishable between the two different models for the calculations from this work and so are not shown, but the $g(r, \Omega_1)$ results are significantly different, and are a graphic illustration of the fact that this is a site-site theory which properly solves the so-called auxiliary site problem.

In calculating results for this model, we use a radial grid of 2048 points and range of 30 Å. For clarity and emphasis, we present the RISM and PISM results in Figs. 5(a)–5(c), and the results using the new closure (here with $\eta=\rho$, the standard construction) in Figs. 6(a)–6(c). The failures of the RISM and PISM equations are evident in this system. The $g_{ij}(r)$ results for the new set of equations show clear, significant improvements in all main features over the standard results, and indeed are all but indistinguishable from the simulation results.

The thermodynamic results for this model are summarized in Tables III and IV, where we label CSL-HNC and mHNC as the standard CSL equations with the HNC-like closure and the molecular equations from this work, respectively. For the Kirkwood-Buff G result, we report the results as calculated from the various possible site origins. The results agree to nearly within numerical error across the various site origins, which is a significant and important result as compared to the CSL-HNC equations. In contrast to the homonuclear models, the results for $\beta U/N$ are essentially equivalent to the CSL-HNC result. It appears to us that the predictions for the thermodynamic quantities for both diagrammatically proper theories are dominated by $g_{\text{ClCl}}(r)$, which is essentially the same in all theories for this particular model. However, these results may prove to be limited by our numerical method, and will be a subject of future investigation.

The angular dependence for this model is summarized in Fig. 7, in which we plot $g(r_{\text{Cl-Cl}}, \theta_1)$. Unlike the homonuclear cases, this model, as expected from the model parameters, is almost completely symmetric, with only the first solvation shell exhibiting any significant asymmetry. We should also point out that the location of the first solvation shell maximum at $\theta_1=0$, not $\pi/2$ as in the homonuclear cases, is consistent with there being little to no direct angular correlation for this model. The inclusion of a H potential in this system is enough to emphasize this dipolelike orientation in the potential, but the H interaction is so small that the intricately detailed effects noted for the angular correlations in the homonuclear cases are not indicated here. A graphic contrast to this model is presented by setting the H potential interaction to 0, which is shown in Fig. 8. For confirmation that this resulting auxiliary site was properly accounted for in the theory, we also calculated the result of changing the bond length as well as zeroing the H potential, and, as expected, the $g_{\text{ClCl}}(r, \theta_1)$ results for any bond length investigated were completely identical.

This figure demonstrates two things. First, because the H potential is nonexistent, the results for this model are a graphic illustration of the fact that this theory is molecularly complete, and thus does not suffer the so-called auxiliary-site problem, in stark contrast to the RISM and PISM theories. By molecularly complete, we mean that each site-site function in the theory contains the same diagrammatic series, diagrammatically exact to order ρ^1 , though each such series is written in a different coordinate system. This is not the case for the RISM and PISM cases, and is the primary reason for the auxiliary site problem in each, as well as the reason that the Kirkwood-Buff G functions are not equivalent in the PISM theory. Furthermore, the generating function of the theory used here is the full molecular f -bond, and thus any expansion element of $\mathbf{c}$ which includes the phantom H–H or H–Cl interaction for this model is simply 0, by construction, and thus trivially propagated by the OZ equations. Second, this figure shows that, while the $g(r)$'s and some thermodynamic quantities may be unaffected by the presence or absence of a small potential around the H, the angular projections and averages can still be significantly and profoundly affected by that potential. This fact will have significant implications in the growing use of the 3D and higher order correlation functions in the study of site-site molecular fluids.

IV. Conclusions

In this work, we have analyzed the molecular pair distribution function for site-site model potentials. We have extended the diagrammatically proper decomposition of Chandler *et al.*¹² to full angular generality for diatomic site-site molecules, and have used this decomposition to resum, or renormalize, the molecular functions in terms of the resulting orientation-dependent site-site functions. A significant formal result of this analysis is the explicit derivation of the radial diagrammatically proper interaction site theory as a minimal angular basis set approximation to the orientation-dependent molecular fluid theory. This analysis of the minimal site-renormalized theory was then inverted to derive the complete molecular theory as generated from a generalization of the PISM-OZ equations to the orientation-dependent Blum-Torruella expansions. Another significant formal result was that certain subclasses of diagrams which would otherwise be considered diagrammatically improper are instead properly included when the complete angular dependence of site-site molecular interactions is formally accounted for. Thus, we have derived the exact orientation-dependent PISM-OZ analog of the molecular OZ theory. We calculated the radial and 3D distribution functions using the molecular HNC approximation and the isotropic basis for four simple diatomic models, and found improved results for all of the models examined, and in the heteronuclear case the quantitative structural improvement was particularly significant. In addition, the thermodynamic averages for the heteronuclear model preserved necessary symmetries across the various site-site origin pairs of the model. This is also a significant improvement over the standard PISM equations.

The completeness of this theory with respect to the molecular potential of the model under consideration was demonstrated through considering the so-called auxiliary site problem. In particular, because the theory here includes the full molecular potential, we were able to address the problem of auxiliary sites by turning off the interaction potential of one of the sites in the heteronuclear model investigated. The radially symmetric distribution functions and site-site pair dependent thermodynamic quantities investigated here were left unaffected by this for the theory developed here, but the effect of zeroing out the interaction of one of the sites in the angular distribution functions was significant. These results indicate that, first, as would be expected of a molecularly complete theory, the auxiliary site problem that affects the RISM and PISM theory is solved in this case. Second, the introduction of a small potential to such auxiliary sites in molecular potentials in order to avoid this problem in the application of RISM and PISM equations unavoidably introduces nonphysical behavior into higher order correlations which might be generated by or inferred from these theories. This has strong implications for the continuing development of the 3D generalizations of these theories.

These results suggest that further analytic work in this area will be useful. In particular, it seems clear to us that this analysis should be especially useful upon extension to polar fluids. In addition, though the computational method used here may be somewhat cumbersome, the extension to molecules with arbitrary numbers of sites is straightforward. Combined with the fact that the 3D distribution functions are easy to generate from this theory, the work set forth here should provide both a formal and a numerical path forward in investigating more complex molecules.

Acknowledgements

The authors gratefully acknowledge the support of the Robert A. Welch Foundation and NIH. K.M.D. also gratefully acknowledges the partial support of the National Science Foundation under NIRT Award No. 29963, and Professor G. Stell and the SUNY-Stony Brook Chemistry Department for graciously hosting him through the completion of this work. They would also like to thank Dr. M. Marucho for the work and discussions which led them to reexamine their work on the effective density term, η , and to realize that they had used a special case of the more general result.

References

1. Blum L, Torruella AJ. *J Chem Phys* 1971;56:303.
2. Fries PH, Patey GN. *J Chem Phys* 1985;82:429.
3. Carlevaro CM, Blum L, Vericat F. *J Chem Phys* 2003;119:5198.
4. Fries PH, Richardi J. *J Chem Phys* 2000;113:9169.
5. Yoshida N, Kato S. *J Chem Phys* 2000;113:4974.
6. Andersen HC, Chandler D. *J Chem Phys* 1972;57:1918.
7. Chandler, D. *The Liquid State of Matter: Fluids, Simple and Complex*. Montroll, EW.; Lebowitz, JL., editors. North-Holland; Amsterdam: 1982. p. 275
8. Cortis CM, Rossky PJ, Friesner RA. *J Chem Phys* 1997;107:6400.
9. Kovalenko A, Hirata F. *J Chem Phys* 1999;110:10095.
10. Beglov D, Roux B. *J Chem Phys* 1996;104:8678.
11. Chandler D, Andersen HC. *J Chem Phys* 1972;57:1930.
12. Chandler D, Silbey R, Ladanyi BM. *Mol Phys* 1982;46:1335.
13. Rossky PJ, Chiles RA. *Mol Phys* 1984;51:661.
14. Høye JS, Stell G. *J Chem Phys* 1976;65:18.
15. Lue L, Blankschtein D. *J Chem Phys* 1995;102:4203.
16. Hansen, JP.; McDonald, IR. *Theory of Simple Liquids*. 2nd. Academic; London: 1986.
17. Dyer KM, Perkyns JS, Pettitt BM. *J Chem Phys* 2005;123:204512. [PubMed: 16351286]
18. Perkyns JS, Dyer KM, Pettitt BM. *J Chem Phys* 2002;116:9404.
19. Dyer KM, Perkyns JS, Pettitt BM. *J Chem Phys* 2002;116:9413.
20. Rossky PJ, Dale WDT. *J Chem Phys* 1980;73:2457.
21. Gray, CG.; Gubbins, KE. *Theory of Molecular Fluids*. 1. Clarendon; Oxford: 1984.
22. Friedman, HL. *A Course in Statistical Mechanics*. Prentice-Hall; Engle-wood Cliffs, NJ: 1985.
23. Allnatt AR. *Mol Phys* 1964;8:533.
24. Ng KC. *J Chem Phys* 1974;61:2680.
25. Ladanyi BM, Chandler D. *J Chem Phys* 1975;62:4308.
26. Edmonds, AR. *Angular Momenta in Quantum Mechanics*. Princeton University Press; Princeton: 1960.
27. Messiah, A. *Quantum Mechanics*. Wiley; New York: 1958.
28. Wertheim MS. *J Chem Phys* 1987;87:7323.
29. Kalyuzhnyi YV, Stell G. *Mol Phys* 1993;78:1247.
30. Stell G. *Physica A* 1996;231:1.
31. Dyer KM, Perkyns JS, Pettitt BM. *J Chem Phys* 2005;122:236101. [PubMed: 16008496]
32. Marucho M, Pettitt BM. *J Chem Phys* 2007;126:124107. [PubMed: 17411108]
33. Perkyns JS, Pettitt BM. *Theor Chem Acc* 1997;96:61.
34. Hirata F, Pettitt BM, Rossky PJ. *J Chem Phys* 1982;77:509.

FIG. 1.

Numerical results for the site-site radial distribution functions for the homonuclear models, the phase points, and model parameters are detailed in the text. r is in units of σ . (a) The $l^* = 0.329$ model, the solid line is the simulation result, and the dashed line is the result from this work, with $\eta = \rho/2.668$. The dotted line is the result from this work with $\eta = \rho/1.24$. (b) The $l^* = 0.547$ model, the line types are the same as in (a), with $\eta = \rho$ here. (c) The $l^* = 0.729$ model, the line types are the same as in (a), with $\eta = \rho$ here.

FIG. 2.

Two different perspectives of the 2D site-site distribution function $g(r_{N-N}, \theta_1)$ for the N_2 model. The ragged edges are an artifact of the plotting procedure and are not representative of the functions themselves. This is a right-handed coordinate system, with three values for $\theta=0, \pi/2, \pi$ marked on the plot for orientational convenience. In this figure and Figs. 4 and 5, σ , the contact distance, is labeled along the x axis for perspective.

FIG. 3.
Two different perspectives of the 2D site-site distribution function $g(r_{\text{Cl-Cl}}, \theta_1)$ for the Cl_2 model.

FIG. 4. Two different perspectives of the 2D site-site distribution function $g(r_{aa}, \theta_1)$ for the $l^*=0.729$ model.

FIG. 5.

The site-site radial distribution functions as predicted by RISM and standard PISM in comparison to simulation for the HCl model. The model and phase parameters are detailed in the text. r is in angstroms. (a) The Cl-Cl $g(r)$ results, the solid line is simulation, the dotted line is the XRISM result, and the dotted-dashed line is the CSL-HNC result. Here, the XRISM and CSL-HNC results are virtually indistinguishable. (b) The H-Cl $g(r)$ results, the line types are the same as in (a). (c) The H-H $g(r)$ results, the line types are the same as in (a).

FIG. 6.

The site-site radial distribution functions predicted by this work for the HCl model. (a) The Cl–Cl $g(r)$ results, with the solid line the simulation result and the dashed line the result of this work, with $\eta=\rho$. (b) The H–Cl $g(r)$ results, with the line types the same as in (a). (c) The H–H $g(r)$ results, with the line types the same as in (a). The H–H and H–Cl theory results are essentially indistinguishable from simulation.

FIG. 7. Two different perspectives of the 2D Cl-Cl site-site distribution function, $g(r_{\text{Cl-Cl}}, \theta_1)$, for the HCl model.

FIG. 8. Two different perspectives of the 2D Cl-Cl site-site distribution function, $g(r_{\text{Cl-Cl}}, \theta_1)$, for the modified HCl model with the H-H interaction parameters set to 0.

TABLE I

Kirkwood-Buff G from simulation and predicted by the various integral equation theories for the homonuclear models in this work.

l^*	CSL-HNC	η -HNC	mHNC	Simulation
0.547	-1.77	-1.92	-1.72	-1.79
0.729	-1.78	-2.11	-1.83	-1.88

TABLE II

Excess internal energy, $\beta U^{\text{ex}}/N\epsilon$ per site pair, from simulation and predicted by the various integral equation theories for the homonuclear models.

l^*	CSL-HNC	η -HNC	mHNC	Simulation
0.329	...	-3.63	-3.53	-3.50
0.547	-3.52	-3.31	-3.08	-2.99
0.729	-3.05	-2.85	-2.61	-2.64

TABLE III

Kirkwood-Buff G from simulation and predicted termwise by the various integral equation theories for the apolar HCl model in this work.

G	XRISM	CSL-HNC	mHNC	Simulation
G_{HH}	-43.6	-47.4	-44.7	-48
G_{HCl}	-43.6	-49.9	-44.6	-48
G_{ClCl}	-43.6	-44.3	-44.3	-48

TABLE IV

Total $\beta U/N$ from simulation and predicted by the various integral equation theories for the apolar HCl model in this work.

	XRISM	CSL-HNC	mHNC	Simulation
$\beta U/N$	-1.69	-1.73	-1.74	-1.69