JCTC Journal of Chemical Theory and Computation

Solutions of the Optimized Closure Integral Equation Theory: Heteronuclear Polyatomic Fluids

M. Marucho,[†] C. T. Kelley,[‡] and B. Montgomery Pettitt*,[†]

Chemistry Department, University of Houston, Houston, Texas 77204-5003, and Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27695-8205

Received August 13, 2007

Abstract: Recently, we developed a thermodynamically optimized integral equation method which has been successfully tested on both simple and homonuclear diatomic Lennard-Jones fluids [*J. Chem. Phys.* **2007**, *126*, 124107]. The systematic evaluation of correlation functions required by the optimization of the chemical potential has shown a clear need for more efficient algorithms to solve these integral equations. In the present paper we introduce a high-performance algorithm which is found to be faster and more efficient than the direct Picard iteration. Here we have utilized this to solve the aforementioned optimized theory for molecules more complex than those considered previously. We analyzed representative models for heteronuclear diatomic and triatomic polar molecular fluids. We include results for several modified SPC-like models for water, obtaining site−site correlation functions in good agreement with simulation data.

I. Introduction

Solvents play an important role in many physical, biological, and chemical processes. Indeed, they govern conformational stability, binding, and functions of biomolecules as well as the details of chemical reactions.[1] Including such solvent effects via computer simulation techniques can be costly since the resulting systems involve a large number of particles with long-range interactions and demand substantial computing time and memory storage.

Alternatively, a variety of computational schemes have been developed to capture the dominant solvation effects on large solutes using approximate representations of solvents. The simplest proposals, namely the continuum models, neglect many essential correlations in solvents and are thus seriously limited.[2] Integral equation theories (IET) offer a compromise between accuracy and computational expense.[3] They provide a powerful theoretical tool for computing approximate pair distribution functions, $g(r)$, from which it is possible to get structural and thermodynamic properties for fluids.[1,4,5] These theories are usually formulated by providing two set of equations, namely the integral equation and a closure relationship. Beyond obtaining a good qualitative description of such systems, closure approximations borrowed from atomic fluids like Percus−Yevick(PY)[6,7] and hypernetted chain (HNC)[8,9] have not produced uniformly good quantitative results for molecular fluids in general. Whether this is due to the propagating integral equation or the closure is not always clear.[10]

As with methods arising from interaction site model representations of molecular liquids,[10] these approximations generate thermodynamic inconsistency as well as quantitative limitations in the description of the short- and long-range structure.[11−21] Indeed, to find an approximate theory describing these properties as accurately as possible and at low computational cost requires the development of more sophisticated theories. In fact, during the last quarter of the century, a significant number of different approaches have been proposed to overcome those difficulties, including new integral equations and/or new closure approximations.[22−29]

Recently, we have developed a thermodynamically consistent integral equation theory which has been successfully tested on both simple and homonuclear diatomic Lennard-

* Corresponding author e-mail: pettitt@uh.edu.
[†] University of Houston.
[‡] North Carolina State University.

10.1021/ct700202h CCC: $40.75
Published on Web 02/21/2008

Jones fluids.[30] A new closure approximation is obtained by using the Percus' functional expansion to first order in the density as a target functional generator.[6,31,32] It depends on parameters providing a smooth transition interpolating between PY and HNC closures which can be variationally optimized. The chemical potential functional is minimized then to yield the values of the parameters. This optimized closure approximation can be coupled with a variant of the diagrammatically proper integral equation introduced by this laboratory,[29] having a density matrix with nondiagonal elements containing screened densities. The simplicity of the expressions involved in the resulting theory have allowed us to also obtain an approximate analytic expression for the molecular excess chemical potential which is minimized to estimate the numerical value of the free parameters that defines the closure. Indeed, the success of this approach is largely due to the fact that for molecular fluids in general there are regions of the phase diagram where PY and HNC usually bracket the simulated pressure and densities. In such a case, one achieves thermodynamic self-consistency at the free energy optimum.

Given that approach, the purpose of this paper centers on implementing a more efficient numerical approach to solve the aforementioned optimized closure approximation for molecules more complex than the homonuclear diatomic fluids considered previously and in particular water. The paper is organized as follows. In section II, we introduce the approximate theory including the expression for the closure approximation, the integral equation, and the analytic approximation for the molecular excess chemical potential μ_{ex}. In section III, we describe the computational scheme used to solve these equations numerically. We also compare this algorithm with direct Picard iteration and other solvers to compare with the literature.[34−38] In section IV, we present the numerical solution obtained for the pair site−site correlation function of two representative models of heteronuclear Lennard-Jones fluids as well as of that modified SPC models for water. Further, we compare our predictions with the corresponding results of MD simulation. Finally, in section V, we summarize the central finding of the article, leaving the details of our computational scheme for the Appendix.

II. Theory

In our recently introduced optimized integral equation theory for homonuclear fluids[30] we found that the optimal parametrization is not universal. As anticipated, it depends on the thermodynamic state, the internal structure of the molecules, and consequently on the different species conforming the molecules. In this way, the extension of the closure approximation for a diagrammatically proper interaction site model (PISM) representation of heteronuclear molecular fluids[24,26,27,39,40] is easily obtained by assigning species labels to the parametrization performed for one-component fluids. Specifically, it reads

$$c^o_{\alpha\gamma}(r) = -a^o_{\alpha\gamma}e^{-\beta u_{\alpha\gamma}(r)} + (1+a^o_{\alpha\gamma})e^{[-\beta u_{\alpha\gamma}(r)+t_{\alpha\gamma}o(r)/(1+a_{\alpha\gamma}o)]} - 1 - t^o_{\alpha\gamma}(r)$$

$$c^r_{\alpha\gamma}(r) = \frac{(1+a^o_{\alpha\gamma})}{(1+a^r_{\alpha\gamma})}t^r_{\alpha\gamma}(r)e^{[-\beta u_{\alpha\gamma}(r)+t^o_{\alpha\gamma}(r)/(1+a^o_{\alpha\gamma})]} - t^r_{\alpha\gamma}(r)$$

$$c^l_{\alpha\gamma}(r) = \frac{(1+a^o_{\alpha\gamma})}{(1+a^r_{\gamma\alpha})}t^l_{\alpha\gamma}(r)e^{[-\beta u_{\alpha\gamma}(r)+t^o_{\alpha\gamma}(r)/(1+a^o_{\alpha\gamma})]} - t^l_{\alpha\gamma}(r)$$

$$c^b_{\alpha\gamma}(r) = (1+a^o_{\alpha\gamma})\left[\frac{t^b_{\alpha\gamma}(r)}{(1+a_\alpha\gamma^b)} + \frac{t^r_{\alpha\gamma}(r)t^l_{\alpha\gamma}(r)}{(1+a^r_{\alpha\gamma})(1+a^r_{\gamma\alpha})}\right]\times e^{[-\beta u_{\alpha\gamma}(r)+t^o_{\alpha\gamma}(r)/(1+a^o_{\alpha\gamma})]} - t^b_{\alpha\gamma}(r) \quad (1)$$

where $t^i_{\alpha\gamma}(r)$ and $c^i_{\alpha\gamma}(r)$ represent the contribution associated with the ith group of terms to the indirect and direct site−site correlation function between sites α and γ, respectively. Here the superscript represents the proper subclasses ($i = o$, l, r, or b) for none, left, right, and both sets of integrals in standard notation.[22,23] The set of parameters, $a^i_{\alpha\gamma}$, is composed of unknowns to be determined in the optimization. The site−site correlation functions are given by the sum of the four components, for instance for the indirect part

$$t_{\alpha\gamma}(r) = t^o_{\alpha\gamma}(r) + t^r_{\alpha\gamma}(r) + t^l_{\alpha\gamma}(r) + t^b_{\alpha\gamma}(r)$$

and so on for h and c. In the latter expression $h_{\alpha\gamma}(r) = g_{\alpha\gamma}(r) - 1$ represents the total site−site correlation function. This closure becomes PISM-HNC and PISM-PY approximations[24] for $a^o_{\alpha\gamma} = a^r_{\alpha\gamma} = a^b_{\alpha\gamma} = a_{\alpha\gamma}$ and $a_{\alpha\gamma} \to 0$ and $a_{\alpha\gamma} \to \infty$, respectively. Expressions 1 can be seen to be a simple generalization of our approximation for heteronuclear proper site−site molecular fluids. Moreover, in the absence of intramolecular correlations, only the o elements are nonzero, and the closure for $c^o_{\alpha\gamma}(r)$ becomes the result previously obtained for multicomponent atomic fluids. We refer the reader to ref 30 for a detailed description of the previous approach.

This approximate theory is completed by coupling the closure approximation 1 with a recently introduced integral equation[29] which, in Fourier space, reads

$$\hat{\mathbf{H}}(k) = \hat{\mathbf{C}}(k) + [\hat{\mathbf{C}}(k) + \hat{\mathbf{S}}(k)]\hat{\boldsymbol{\rho}}[\hat{\mathbf{H}}(k) + \hat{\mathbf{S}}(k)] \quad (2)$$

where $\hat{\mathbf{H}}(k)$ and $\hat{\mathbf{C}}(k)$ are the Fourier transform of $\mathbf{H}(r)$ and $\mathbf{C}(r)$, respectively. Each of the correlation functions appearing above is a symmetric matrix in the form

$$\mathbf{Q}(k) = \begin{pmatrix} \mathbf{Q}_{11}(k) & \mathbf{Q}_{12}(k) \\ \mathbf{Q}_{21}(k) & \mathbf{Q}_{22}(k) \end{pmatrix}, \quad \mathbf{Q}_{\alpha\gamma}(k) = \begin{pmatrix} q^o_{\alpha\gamma}(k) & q^r_{\alpha\gamma}(k) \\ q^l_{\alpha\gamma}(k) & q^b_{\alpha\gamma}(k) \end{pmatrix}$$

where $\mathbf{Q}$ represents $\mathbf{H}$ or $\mathbf{C}$. The screened density and the renormalized intramolecular matrices $\bar{\boldsymbol{\rho}}$ and $\hat{\mathbf{S}}(k)$ for triatomic heteronuclear molecules are given by the following expressions

$$\bar{\boldsymbol{\rho}} = \begin{pmatrix} \rho & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \rho \end{pmatrix}, \quad \rho = \begin{pmatrix} \rho & \eta \\ \eta & 0 \end{pmatrix}, \quad \hat{\mathbf{S}}(k) = \begin{pmatrix} 0 & \hat{\mathbf{s}}_{12}(k) & \hat{\mathbf{s}}_{13}(k) \\ \hat{\mathbf{s}}_{21}(k) & 0 & \hat{\mathbf{s}}_{23}(k) \\ \hat{\mathbf{s}}_{31}(k) & \hat{\mathbf{s}}_{32}(k) & 0 \end{pmatrix},$$

$$\hat{\mathbf{s}}_{ij}(k) = \begin{pmatrix} 0 & 0 \\ 0 & (\sin(kL_{ij})/\eta kL_{ij}) \end{pmatrix} \quad (3)$$

with η being the screened density whose approximation to

first order[29] can be numerically computed by the following expression

$$\eta = \rho/[1 + \rho \int d\mathbf{r}_2\{f_{21}(r_{12})f_{11}(r_2) + f_{23}(r_{12})f_{31}(r_2) + f_{22}(r_{12})f_{21}(r_2)\}|_{r_1=L_{12}} + \rho\int d\mathbf{r}_2\{f_{33}(r_{12})f_{31}(r_2) + f_{32}(r_{12})f_{21}(r_2) + f_{31}(r_{12})f_{11}(r_2)\}|_{r_1=L_{13}}] \quad (4)$$

In the latter formula, ρ represents the molecular density, $L_{\alpha\gamma}$ is the bond length between the two sites α and γ, and $f_{\alpha\gamma}$ is the corresponding Mayer function.

The simplicity of the previous expressions has allowed us to also obtain an approximate analytic expression for the molecular excess chemical potential μ_{ex}. This is convenient; much as with density functional theories this is the quantity which is minimized to estimate the value of the free parameters that define the closure approximation 1. In fact, it is obtained by a straightforward extension of the calculation carried out in section IV of ref 30. The resulting expression reads

$$\beta\mu_{ex}(\{a^i_{ab}\}) \simeq -\rho \sum_{\alpha\gamma} \int d\mathbf{r}\Bigg\{ h_{\alpha\gamma}(r) - \mathscr{T}[t^o_{\alpha\gamma}(r),h^o_{\alpha\gamma}(r),a^o_{\alpha\gamma}] \left[t^o_{\alpha\gamma}(r) + (1 + a^o_{\alpha\gamma})\left[\frac{t^r_{\alpha\gamma}(r)}{(1 + a^r_{\alpha\gamma})} + \frac{t^l_{\alpha\gamma}(r)}{(1 + a^r_{\gamma\alpha})}\right] + \frac{(1 + a^o_{\alpha\gamma})t^b_{\alpha\gamma}(r)}{(1 + a^b_{\alpha\gamma})}\right] - \frac{h^r_{\alpha\gamma}(r)t^l_{\alpha\gamma}(r)}{(1 + a^r_{\gamma\alpha})} - \frac{[h^r_{\alpha\gamma}(r) + h^l_{\alpha\gamma}(r) + h^b_{\alpha\gamma}(r)]t^o_{\alpha\gamma}(r)}{2(1 + a^o_{\alpha\gamma})} \quad (5)$$

where $\mathscr{T}[u, v, w]$ is given by the following formula (eq 12 in ref 30)

$$\mathscr{T}[u, v, w] = \frac{\left\{[v + 1]\ln\left[\frac{y}{w}\right] + \ln[w] + \frac{v(w + 1)}{u}\mathrm{Re}\left[Li_2\left(\frac{y}{w} + 1\right) - Li_2\left(\frac{w + 1}{w}\right)\right]\right\}}{u} \quad (6)$$

with $Li_2(x)$ being the well-known dilogarithm function[41,42] and $y = -w + (w + 1)\exp[u/(w + 1)]$.

The aforementioned optimization procedure certainly requires a fast and efficient computational scheme to evaluate the correlation functions appearing in expression 5. This topic is analyzed in the next section, where a high-performance numerical algorithm is presented.

III. Numerical Analysis

As shown initially by Gillan for atomic fluids,[43] the multiresolution representation of correlation functions into a "coarse'' part, which is expanded in a set of bases functions with coefficients determined by the Newton-Rapshon (NR) scheme, and the "fine'' part, which is evaluated numerically by direct Picard iteration, leads to a significant reduction in the computational complexity and cost to solve liquid state integral equation theories. Subsequently modified and adapted for more complex fluids, this method has provided excellent results in general.[44−46] In fact, it is well-known that a fairly rapidly convergence to a solution is achieved when these iterative solvers are started from a sufficiently good initial guess. Further, the choice of the basis functions employed to expand the correlation functions and the method utilized to determine the corresponding weight of their projections play a fundamental role in the rate of convergence. These methods, however, require the analytical expression for the Jacobian matrix to be computationally most efficient.

Based on these considerations, the numerical method proposed in the present study consists of expanding the indirect correlation function on a sine function basis set, a natural expansion within a scheme that already involves a systematic evaluation of sine Fourier transforms. Indeed, unidimensional transforms appear in the integral eq 2 using the angular average of the spherically symmetric correlation functions defined in the three-dimensional Fourier transforms. As shown in ref 45, the so named "unidimensional'' correlation functions

$$\mathscr{T}_{\alpha\gamma}(r) \equiv rt_{\alpha\gamma}(r), \quad \mathscr{C}_{\alpha\gamma}(r) \equiv rc_{\alpha\gamma}(r), \quad \alpha, \gamma = 1, .., 2m \quad (7)$$

satisfy the following properties

$$\mathscr{A}_{\alpha\gamma}(r) = \frac{1}{2\pi^2}\int_0^\infty dk \sin(kr)\hat{\mathscr{A}}_{\alpha\gamma}(k) \quad (8)$$

$$\hat{\mathscr{A}}_{\alpha\gamma}(k) = 4\pi\int_0^\infty dk \sin(kr)\mathscr{A}_{\alpha\gamma}(r) \quad (9)$$

where $\mathscr{A}$ represents $\mathscr{C}$ or $\mathscr{T}$. Consequently, eqs 1 and 2 become

$$\mathscr{C}^o_{\alpha\gamma}(r) = -ra^o_{\alpha\gamma}e^{-\beta u_{\alpha\gamma}(r)} + r(1 + a^o_{\alpha\gamma})e^{[-\beta u_{\alpha\gamma}(r) + \mathscr{T}_{\alpha\gamma o}(r)/[r(1+a_{\alpha\gamma o})]]} - r - \mathscr{T}^o_{\alpha\gamma}(r)$$

$$\mathscr{C}^r_{\alpha\gamma}(r) = \frac{(1 + a^o_{\alpha\gamma})}{(1 + a^r_{\alpha\gamma})}\mathscr{T}^r_{\alpha\gamma}(r)e^{[-\beta u_{\alpha\gamma}(r) + \mathscr{T}_{\alpha\gamma}{}^o(r)/[r(1+a_{\alpha\gamma o})]]} - \mathscr{T}^r_{\alpha\gamma}(r)$$

$$\mathscr{C}^l_{\alpha\gamma}(r) = \frac{(1 + a^o_{\alpha\gamma})}{(1 + a^r_{\gamma\alpha})}\mathscr{T}^l_{\alpha\gamma}(r)e^{[-\beta u_{\alpha\gamma}(r) + \mathscr{T}_{\alpha\gamma o}(r)/[r(1+a_{\alpha\gamma o})]]} - \mathscr{T}^l_{\alpha\gamma}(r)$$

$$\mathscr{C}^b_{\alpha\gamma}(r) = (1 + a^o_{\alpha\gamma})\left[\frac{\mathscr{T}^b_{\alpha\gamma}(r)}{r(1 + a^b_{\alpha\gamma})} + \frac{\mathscr{T}^r_{\alpha\gamma}(r)\mathscr{T}^l_{\alpha g}(r)}{r^2(1 + a^r_{\alpha\gamma})(1 + a^r_{\gamma\alpha})}\right] \times e^{[-\beta u_{\alpha\gamma}(r) + \mathscr{T}^o_{\alpha\gamma}(r)/[r(1+a^o_{\alpha\gamma})]]} - \mathscr{T}^b_{\alpha\gamma}(r) \quad (10)$$

and

$$\hat{\mathscr{T}}(k) = k[k\bar{\boldsymbol{\rho}}^{-1}[k\bar{\boldsymbol{\rho}}^{-1} - \hat{\mathscr{C}}(k) - k\hat{\mathbf{S}}(k)]^{-1}\bar{\boldsymbol{\rho}}^{-1} - \bar{\boldsymbol{\rho}}^{-1}] - k\hat{\mathbf{S}}(k) - \hat{\mathscr{C}}(k) \equiv IE(\hat{\mathscr{C}}(k)) \quad (11)$$

respectively. Clearly, the discretization of eq 8 yields the

following definition of the coarse and fine part of the indirect correlation function

$$\mathcal{T}_{\alpha\gamma,j} = \frac{\Delta k}{2\pi^2}\sum_{n=1}^{N-1}\sin\left(\frac{jn\pi}{N}\right)\hat{\mathcal{T}}_{\alpha\gamma,n} = \frac{\Delta k}{2\pi^2}\sum_{n=1}^{M}\sin\left(\frac{jn\pi}{N}\right)\hat{\mathcal{T}}_{\alpha\gamma,n} + \frac{\Delta k}{2\pi^2}\sum_{n=M+1}^{N-1}\sin\left(\frac{jn\pi}{N}\right)\hat{\mathcal{T}}_{\alpha\gamma,n}$$

$$\equiv \text{coarse part} + \text{fine part} \quad (12)$$

The latter equation is valid for $j = 1, .., N-1$ and $\alpha, \gamma = 1, .., 2m$, being $\mathcal{A}_{\alpha\gamma}(r_j) \equiv \mathcal{A}_{\alpha\gamma,j}$, Δr and $\Delta k = \pi/(\Delta r N)$ the corresponding mesh size in distance and reciprocal space, respectively, N is the number of points on the grid, and M is an integer to be fixed later. Obviously, the coefficients for these expansions are directly the sine Fourier components, which can be rapidly evaluated, when required, via Fast Fourier Transform (FFT) techniques. In fact, the discretization of eq 9 yields

$$\hat{\mathcal{A}}_{\alpha\gamma,n} \equiv 4\pi\Delta r\sum_{j=1}^{N-1}\sin\left(\frac{jn\pi}{N}\right)\mathcal{A}_{\alpha\gamma,j}, \quad \alpha, \gamma = 1, .., 2m, \quad n = 1..N-1 \quad (13)$$

In this way, the numerical scheme deriving from the decomposition 12 has two components. First, a fast and robust Newton-GMRES algorithm,[47] which approximates the solution for the weight of the projections in each iteration, is initially utilized to solve the set of nonlinear eqs 10 and 11 for the first M sine Fourier components $\{\hat{\mathcal{T}}_{\alpha\gamma,n}\}$ (primary contribution), keeping the remaining components fixed. This is followed by one direct Picard iteration to refine the higher sine Fourier components, namely for $n = M+1,..,$ $N-1$. This sequence is repeated to get convergence. Other variants of GMRES have been used before in this area.[33−38]

The first phase involves the solution of the system of equations

$$R_{\alpha\gamma,n} \equiv \hat{\mathcal{T}}_{\alpha\gamma,n} - [IE(\hat{\mathcal{C}})]_{\alpha\gamma,n} = 0, \quad \alpha, \gamma = 1,..,2m \quad (14)$$

with respect to the first M sine Fourier components $\hat{\mathcal{T}}_{\alpha\gamma,n}$ ($n = 1,..M$). This requires the evaluation of the Jacobian matrix

$$J_{\alpha\gamma,n;\beta\zeta,p} = \frac{\partial R_{\alpha\gamma,n}}{\partial\hat{\mathcal{T}}_{\beta\zeta,p}} = \delta_{\alpha\beta}\delta_{\gamma\zeta}\delta_{np} - \frac{\partial[IE(\hat{\mathcal{C}})]_{\alpha\gamma,n}}{\partial\hat{\mathcal{T}}_{\beta\zeta,p}}, \quad n, p = 1,..,M, \quad \alpha, \gamma = 1, .., 2m \quad (15)$$

in which the derivative appearing on the right side of the latter equation is computed by the chain rule. Indeed, each sine Fourier component n of the element α, γ of the direct correlation function, $\hat{\mathcal{C}}_{\alpha\gamma,n}$, depends on the corresponding direct correlation function defined over the entire grid as given by eq 13, and each one of those components, namely $\mathcal{C}_{\alpha\gamma,j}$, depends on elements of the indirect correlation function via the closure relationships 10. Note that the latter relationships can be conveniently written in terms of correlation function matrices of dimensions $[2m \times 2m]$ like those appearing in the integral eq 11. Using the following identities between the different matrix notations: $\mathcal{A}^{o}_{\alpha\gamma}(r) = \mathcal{A}_{2\alpha-1,2\gamma-1}(r)$, $\mathcal{A}^{r}_{\alpha\gamma}(r) = \mathcal{A}_{2\alpha-1,2\gamma}(r)$, $\mathcal{A}^{l}_{\alpha\gamma}(r) = \mathcal{A}_{2\alpha,2\gamma-1}(r)$, and $\mathcal{A}_{2\alpha,2\gamma}(r) = \mathcal{A}_{2\alpha,2\gamma}(r)$ for $\alpha, \gamma = 1, .., m$, we can formally express the four discretized closure approximations 10 by the following general relationship

$$\mathcal{C}_{2\alpha-i,2\gamma-j,p} = CL^{ij}\{\mathcal{T}_{2\alpha-1,2\gamma-1,p}, \mathcal{T}_{2\alpha-1,2\gamma,p}, \mathcal{T}_{2\alpha,2\gamma-1,p}, \mathcal{T}_{2\alpha,2\gamma,p}\}, \quad \alpha, \gamma = 1, .., m, \quad i, j = 0, 1 \quad (16)$$

the set of pairs $(i,j) = (0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ representing the subclasses *o*, *r*, *l*, and *b*, respectively.

Eq 16 clearly shows that the analytic expression for $\partial\mathcal{C}_{\beta\zeta,j}/\partial\mathcal{T}_{\alpha\gamma,j}$ contains new nonzero elements and new contributions for those found in other IETs requiring only one of these relationships (subclass *o* for instance) to close the theory and having the correspondence between each element of the matrices $\mathcal{C}$ and $\mathcal{T}$. For instance, the derivative $\partial\mathcal{C}_{\beta\zeta,j}/\partial\mathcal{T}_{11,j}$ is nonzero for the following pair of elements $(\beta, \zeta) = (1,1)$, $(1,2)$, $(2,1)$, and $(2,2)$. Further, each is different from the other. Finally, each one of the elements $\mathcal{T}_{\alpha\gamma,j}$ appearing in eq 16 depends on the sine Fourier components of the indirect correlation function defined over the entire grid in the reciprocal space as shown by eq 12. As a result of these complicated relationships, the chain rule yields the following analytical expression for the Jacobian matrix

$$J_{\alpha\gamma,n;\beta\zeta,p} = \delta_{\alpha\beta}\delta_{\gamma}\zeta_{\delta np} - \sum_{\mu,\nu=1}^{2m}\{-\delta_{\alpha\mu}\delta_{\gamma\nu} + [\mathbf{I} + [\hat{\mathbf{H}} + \hat{\mathbf{S}}]\bar{\boldsymbol{\rho}}]_{\alpha\mu,n}[\mathbf{I} + \bar{\boldsymbol{\rho}}[\hat{\mathbf{H}} + \hat{\mathbf{S}}]]_{\nu\gamma,n}\}\Phi_{\mu\nu,n;\beta\zeta,p} \quad (17)$$

in which

$$\Phi_{\mu\nu,n;\beta\zeta,p} \equiv \frac{2}{N}\sum_{j=1}^{N-1}\sin\left(\frac{jn\pi}{N}\right)\sin\left(\frac{jp\pi}{N}\right)\frac{\partial\mathcal{C}_{\mu\nu,j}}{\partial\mathcal{T}_{\beta\zeta,j}} \quad (18)$$

Note that the expression for $\partial\mathcal{C}_{\mu\nu,j}/\partial\mathcal{T}_{\beta\zeta,j}$ is easily obtained from eqs 16 and 10 and that the coefficients 18 are numerically computed using FFTs. The computational cost of computing the full Jacobian and the comparison to the cost of a matrix-free approach are left for a future discussion.

Finally, the analytical expressions for the Jacobian and residual are utilized in the aforementioned nonlinear solve as follows. We provide an initial guess for $\mathcal{T}^{guess}_{\alpha\gamma,j}$ for $j = 1, .., N-1$, and $\alpha, \gamma = 1, .., 2m$, from where we evaluate the first M sine Fourier components $\hat{\mathcal{T}}^{guess}_{\alpha\gamma,k}$ using FFT (eq 13). We also evaluate the elements $\mathcal{C}^{guess}_{\alpha\gamma,j}$ defined by the closure relationship 16 and subsequently the entire set of sine Fourier components for $\hat{\mathcal{C}}^{guess}_{\alpha\gamma,n}$ via FFT (eq 13). These elements are required to evaluate the expressions given for the residual (14) and the Jacobian (17) matrices. Then, the new estimates of the first M components $\hat{\mathcal{T}}^{new}_{\alpha\gamma,k}$ ($k = 1, .., M$) are obtained by solving the set of linear equations

$$\sum_{\mu,\nu=1}^{2m}\sum_{p=1}^{M}J^{guess}_{\alpha\gamma,n;\mu\nu,p}\Delta\hat{\mathcal{T}}_{\mu\nu,p} = R^{guess}_{\alpha\gamma,n}, \quad \alpha, \gamma = 1, .., 2m, \quad n = 1, .., M \quad (19)$$

for $\Delta\hat{\mathcal{T}}_{\alpha\gamma,k} \equiv \hat{\mathcal{T}}^{new}_{\alpha\gamma,k} - \hat{\mathcal{T}}^{guess}_{\alpha\gamma,k}$. To achieve this, we use GMRES. We repeat the previous calculation iteratively to

get the desired convergence. The first phase is accomplished by choosing a positive value for the parameter δ such that

$$\sqrt{\sum_{\mu,\nu=1}^{2m}\sum_{p=1}^{M}[\Delta\hat{\mathcal{T}}_{\mu\nu,p}]^2} < \sqrt{M}\delta \qquad (20)$$

The second phase is subsequently implemented to minimize the norm of the matrix 14 with respect to the remaining elements $\hat{\mathcal{T}}_{\alpha\gamma,k}$ ($k = 1 + M$, .., $N - 1$). The new estimates for the higher components are easily obtained by direct evaluation of eqs 10, 13, and 11 as performed by one Picard iteration. The input for this iteration is given by the indirect correlation function coming from eq 12 in which the first M components are those obtained in the first phase and the remaining are those initially kept fixed.

Last, we check the convergence for the entire cycle on a distancelike norm. We use the new sine Fourier components previously obtained in the first and second phases to calculate new estimates for the indirect correlation function, $\mathcal{T}^{new}_{\mu\nu,p}$ ($p = 1$, .., $N - 1$) as given by eq 12. We will obtain the solution at the required precision η when the following condition

$$\sqrt{\sum_{\mu,\nu=1}^{2m}\sum_{p=1}^{N-1}[\mathcal{T}^{new}_{\mu\nu,p} - \mathcal{T}^{guess}_{\mu\nu,p}]^2} < 2m\sqrt{(N-1)\eta} \qquad (21)$$

is fulfilled. Otherwise, we redefine $\mathcal{T}^{guess}_{\mu\nu,p} = \mathcal{T}^{new}_{\mu\nu,p}$ and go back to the first phase.

In considering the global convergence of this proposal, it is well-known that numerical solvers based on an NR scheme have no guarantees of convergence. Rapid local convergence is reachable when a guess sufficiently close to the solution is provided. To achieve this, a nested algorithm is implemented, which approximately solves the nonlinear equation systems 10 and 11 using the above approach on a sequence of meshes, ending with a solution at the target or finest mesh.[47]

We obtain a solution initially at low resolution for a large mesh size $\Delta\xi_o$ and with few points $N_{\xi o}$ such that $\Delta\xi_o(N_{\xi o} - 1) = L$. The convergence is fast and efficient since the complexity and dimensionality of the problem has been significantly reduced and the results are not required to be as accurate as that for the finest mesh. At subsequent levels of iterations, such complexity is gradually increased without affecting the rate of convergence considerably. The reason is the fact that a very good initial guess is obtained at each level (except for the first nested iteration) by using an interpolating linear polynomial splines on the nodes of the solution generated in the previous level. As the number of points N increases, the error in using the aforementioned interpolating polynomial as an approximation to the desired solution tends to zero like N^{-2}. Thus, higher levels in the nested iteration provides closer initial guesses for the iterative solver described above. This guarantees a rapid local convergence without requiring a larger set of basis functions. Consequently, it provides a fast and efficient scheme by which a sufficiently good initial guess to the solution required for the target grid is reached at very low computational cost. For convenience, we used in the present paper a nested grid defined by $\Delta\xi_j = 2^{-j}\Delta\xi_o$ and $N_j - 1 = L/\Delta\xi_j$ with $j = 0$, .., j_{max}, in such a way that the mesh size is reduced by a factor of 2 at each level. Accordingly, the guess for the jth nested level ($j > 0$) is transferred from the solution obtained in the previous level as follows: $\mathcal{T}^{guess}_{\mu\nu,2p} = \mathcal{T}^{solution}_{\mu\nu,p}$ and $\mathcal{T}^{guess}_{\mu\nu,2p+1} = 1/2(\mathcal{T}^{solution}_{\mu\nu,p+1} + \mathcal{T}^{solution}_{\mu\nu,p})$ for $p = 0$, .., $(N_{j-1} - 1)$ and μ, $\nu = 1$, .., $2m$, where $\mathcal{T}_{\mu\nu,0} \equiv 0$ and we use $\mathcal{T}^{guess}_{\mu\nu,2N-1} \cong \mathcal{T}^{solution}_{\mu\nu,N-1}$.

Table 1. Nonpolar Hydrogen Chlorine Fluid: Computing Times (Expressed in Seconds) Obtained on a Common PC Using the Nested Picard Iteration NPI ($M = 0$,$j_{max} = 6$), Our Proposal as Described in Section III OP ($M \approx 50$, $j_{max} = 6$), Plain Picard Iteration PPI ($M = 0$, $j_{max} = 0$), and Our Proposal without the Nested Iteration OP ($M \approx 50$, $j_{max} = 0$)[a]

	HCL[b]							
	$j_{max} = 6$[c]				$j_{max} = 0$ (directly on the target grid)[c]			
	NPI[d]			OP[d]	PPI[d]			OP2[d]
α	0.75	0.65	0.45		0.75	0.65	0.45	
OPT	19.0	20.9	29.0	4.1	66.2	46.5	65.18	10.0
PY	14.84	18.3	22.6	4.2	32.6	28.4	40.4	10.5
HNC	29.7	5.8	11.1	4.4	55.4	13.7	55.6	7.6
CL1	5.2	5.8	8.4	4.55	34.6	16.6	18.3	13.3
CL2	14.3	16.6	23.8	4.3	25.3	30.9	39.9	7.5
CL3	14.7	17.0	24.4	3.9	27.9	28.1	44.1	6.9
CL4	14.9	17.2	25.1	4.7	26.9	27.8	43.2	10.6
CL5	15.0	17.2	24.8	5.5	24.7	28.0	50.2	9.2
CL6	16.5	14.8	21.6	4.7	35.3	44.9	73.4	9.2
CL7	15.0	17.4	25.3	4.5	31.3	34.5	134.8	10.4
CL8	32.1	21.9	28.8	6.2	63.9	45.7	56.2	9.8
CL9	17.0	19.5	28.2	4.0	30.5	36.5	52.3	9.9

[a] α represents the mixing parameter, being the minimum relaxation for α→1. The set of parameters $\{a^i_{\mu\nu}\}$ defines the closure approximation. They are written in the same order as the one presented in section IV. We named OPT = $\{a^i_{\mu\nu}$ = optimized parameters$\}$, PY = $\{a^i_{\mu\nu} = 100\}$, HNC = $\{a^i_{\mu\nu} = 0\}$, CL1 = $\{a^i_{\mu\nu} = 1\}$, CL2 = $\{a^i_{\mu\nu} = 5\}$, CL3 = $\{a^i_{\mu\nu} = 10\}$, CL4 = $\{a^i_{\mu\nu} = 25\}$, CL5 = $\{a^i_{\mu\nu} = 50\}$, CL6 = {2,2,1,4,2,2,3,3,1,4}, CL7 = {4,3,3,3,1,2,2,4,3,9}, CL8 = {13,11,11,9,9,9,6,6,13,9}, and CL9 = {23,21,31,33,20,26,24,24,23,29}. [b] Molecule. [c] Nested levels. [d] Algorithm.

Just for the purpose of comparing relative computing times between several numerical solvers (including the purely nested Picard iteration ($M = 0$) and the present proposal ($M \neq 0$)) we present in Tables 1 and 2 the numerical results obtained for different closure approximations and thermodynamic states for HCl and water (see the next section for details about the feature of these fluids). We have tabulated in columns named NPI, OP, PPI, and OP2 the results obtained on a common PC using the nested Picard iteration ($M = 0$, $j_{max} > 0$), our proposal as described above ($M > 0$, $j_{max} > 0$), plain Picard iteration ($M = 0$, $j_{max} = 0$), and our proposal without the nested iteration ($M > 0$, $j_{max} = 0$), respectively. The results were obtained in the interval [0, $L = 35.84$ Å] on a target grid of mesh size $\Delta r = 0.004375$ Å and number of points $N = 8193$, with a precision parameter η set to 10^{-10}. We initially obtained the approximate solution of the nonlinear equation systems on the coarsest grid of mesh size $\Delta\xi_o = 0.28$ Å and number of points $N_o = 129$, with a precision required of only one significant digit ($\eta_o =$

Table 2. Nonpolar Waterlike Model[a]

	H_2O[b]							
	$j_{max} = 6$[c]				$j_{max} = 0$ (directly on the target grid)[c]			
	NPI[d]			OP[d]	PP[d]			OP2[d]
α	0.75	0.65	0.45		0.75	0.65	0.45	
OPT	130.4	150.3	217.8	13.8	198.4	228.4	324.2	25.7
PY	219.5	253.1	364.5	13.9	328.8	378.7	534.0	23.1
HNC	102.0	118.8	171.0	13.2	157.3	180.5	255.6	22.9
CL1	162.1	186.5	269.9	14.7	210.2	242.6	343.5	24.5
CL2	215.6	249.2	361.3	14.0	278.8	319.0	452.7	24.0
CL3	232.9	269.9	390.8	14.0	301.9	345.1	488.9	24.0
CL4	247.5	286.6	413.0	13.9	371.7	365.5	519.5	23.3
CL5	252.5	292.1	422.9	13.9	325.3	373.8	529.6	23.3
CL6	185.1	213.2	309.3	13.9	240.0	275.8	389.2	23.6
CL7	199.4	229.1	333.2	16.9	257.5	296.5	420.2	25.8
CL8	229.2	275.2	384.6	15.4	301.6	346.8	491.2	23.7
CL9	244.9	283.0	411.1	14.8	316.1	363.1	512.5	23.5

[a] The notation is the same as that in Table 1. [b] Molecule. [c] Nested levels. [d] Algorithm.

10^{-1}). At intermediate levels, this precision was increased gradually. This means that just a few levels ($j_{\max} = 6$) were required in the nested iteration to obtain a very good initial guess for the target grid.

The resulting convergence in the nested Picard iteration was found to depend notably on the value of the parameters defining the closure approximation and on the relaxation parameter. In contrast with our proposal, we had to restart Picard several times on the coarsest grid, adjusting the relaxation parameter to get convergence for many closure approximations. Specifically, we shifted slightly the value of the relaxation parameter and restarted the iteration every time that the rate between residuals of consecutive iterations increases by 30 times or more. In other cases this was not necessary, but the rate of convergence for different closure approximations obtained from the same value of the relaxation parameter was very slow. These facts show a clear restriction in using a purely Picard iteration in optimization programs which demand the fast evaluation of correlation functions at different closure approximations.

In contrast, our algorithm was shown to be faster and more efficient than direct Picard iteration. The reason presumably lies in the fact that only the first $M \sim 0.01N$ sine Fourier components were sufficient to get a good representation of the indirect correlation functions. Clearly, a very high value for M would make inefficient this approach since it gives an enormous system of linear equations. On the other hand, a very small value for M would not dramatically improve the computing times obtained by a direct Picard iteration. We found particularly useful a fixed value for $M \cong 50$ to get the optimal rate of convergence of the present computational scheme for the models analyzed in this article. The parameter δ was updated after each global iteration (second phase) in such a way that $\delta_i = a_i/10$, being that a_i is the error generated for the indirect correlation function after performing i global iterations as given by the left term in expression 21. This algorithm is summarized in the Appendix.

Table 3. Performance of the Nested Iteration Obtained in Our Proposal for Nonpolar Fluids[a]

			molecule			
			HCL		*H_2O*	
nested levels	GRID	DP	GI	RC	GI	RC
$j = 0$	129	1	4	0.25	5	0.2
$j = 1$	257	2	7	0.33	8	0.33
$j = 2$	513	3	10	0.33	11	0.33
$j = 3$	1025	4	13	0.33	14	0.33
$j = 4$	2049	5	16	0.33	17	0.33
$j = 5$	4097	7	19	0.67	20	0.67
$j = 6$	8193	10	23	0.75	24	0.75

[a] We presented the digits of precision (DP) obtained for the residual at each level of the nested iteration as well as the number of global iterations (GI) performed in the corresponding levels. We also presented the rate of convergence (RC = ΔDP/ΔGI) achieved between consecutive levels. These results correspond to the optimized closure approximation.

On the other hand, performing matrix operations and computing Fourier transforms directly on the target grid increased significantly the computing times in both algorithms as shown in columns PPI and OP2. These algorithms also became more unstable and dependent on the initial guess. By comparing the columns NPI and PPI for a Picard iteration and the columns OP and OP2 for our proposal, we conclude that the nested iteration sped up the rate of convergence in both approaches for all the cases analyzed in this work. It is worth mentioning that the results tabulated in the OP2 columns provide an estimation of the computing times that are obtained by numerical solvers based on Gillian/GMRES algorithms.[33−38]

To illustrate the importance of the nested iteration in the performance of our proposal we tabulated in Table 3 the number of global iterations (GI) and the corresponding number of digits of precision (DP) achieved for the residual at each level j of the nested iteration. We also tabulated the rate of convergence (RC = ΔDP/ΔGI) between consecutive levels. Since we have obtained similar results for all the closures previously analyzed in Tables 1 and 2, we only presented in Table 3 those corresponding to the optimized closure approximation. As expected, the results tabulated in column RC indicate that the increase in the dimensionality and complexity of the set of nonlinear equations does not reduce the rate of convergence. Thus, we reached a sufficiently good initial guess for the target grid at a very low computational cost from where the set of nonlinear equations was solved by performing only 4 global iterations.

In the next section, representative models of heteronuclear molecular fluids are analyzed using this computational scheme. To determine the accuracy and efficiency of the theory in predicting the structure of such systems, pair site−site correlation functions are quantitatively compared against MD simulation. The results presented below were obtained with a finest grid of 4097 points and mesh size of 0.00875 Å.

IV. Results and Discussions

A. Diatomics. As a preliminary test, we have numerically solved the equations for two quite different heteronuclear

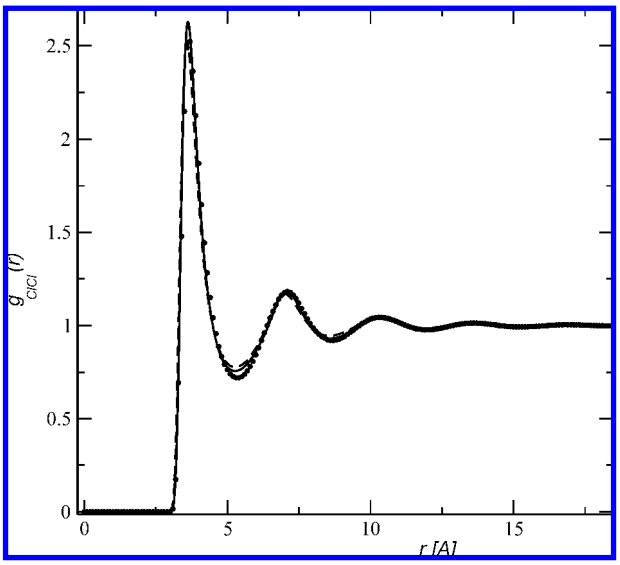

Figure 1. Nonpolar hydrogen chloride: Site−site pair correlation function $g_{ClCl}(r)$ at density $\rho = 0.018$ Å^{-3}, screened density $\eta = \rho/1.361$, and temperature $T = 210$ K. Dashed and solid lines correspond to the PISM-HNC prediction[24] and this work, respectively. Circles correspond to data simulation.

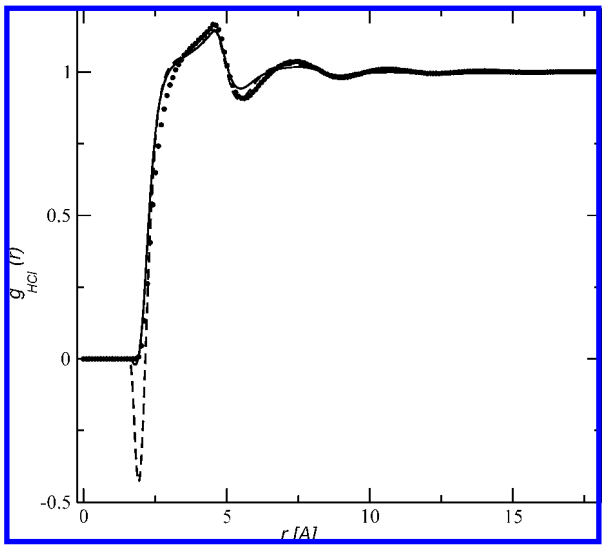

Figure 2. Nonpolar hydrogen chloride: Site−site pair correlation function $g_{HCl}(r)$ at density $\rho = 0.018$ Å^{-3}, screened density $\eta = \rho/1.361$, and temperature $T = 210$ K. Dashed and solid lines correspond to the PISM-HNC prediction[24] and this work, respectively. Circles correspond to data simulation.

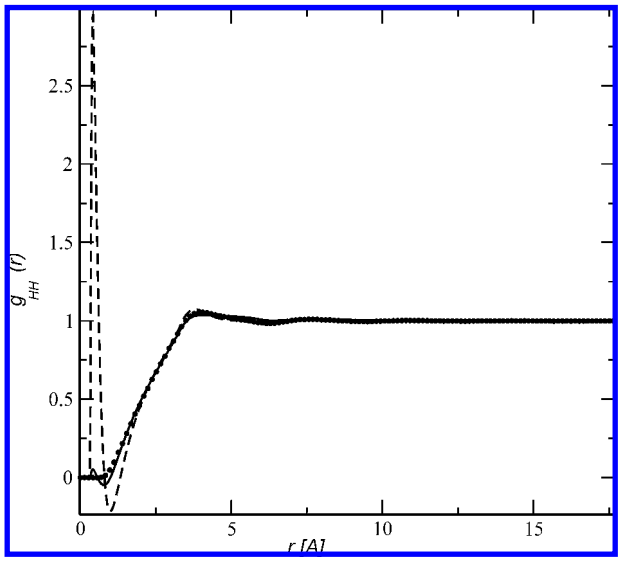

Figure 3. Nonpolar hydrogen chloride: Site−site pair correlation function $g_{HH}(r)$ at density $\rho = 0.018$ Å^{-3}, screened density $\eta = \rho/1.361$, and temperature $T = 210$ K. Dashed and solid lines correspond to the PISM-HNC prediction[24] and this work, respectively. Circles correspond to data simulation.

molecular fluids. The first system corresponds to a Lennard-Jones fluid corresponding to a nonpolar hydrogen-chloride liquid (*HCl*).[27] This is a severe test due to the chemical fact that the hydrogen is completely enveloped in the van der Waals sphere of the chlorine. The predicted site−site pair correlation functions, $g_{ClCl}(r)$, $g_{HCl}(r)$, and $g_{HH}(r)$, at density $\rho = 0.018$ Å^{-3}, screened density $\eta = \rho/1.361$, temperature $T = 210$ K, and bond length $L_{HCl} = 1.3$ Å are plotted in Figures 1−3, respectively. The solid and dashed lines correspond to our prediction and the PISM-HNC approximate theory, respectively, and circle symbols correspond to MD simulation. The parameters which describe the Lennard-Jones potential are $\sigma_{HH} = 0.4$ Å, $\epsilon_{HH}/k_B = 20$ K, $\sigma_{ClCl} = 3.353$ Å, and $\epsilon_{ClCl}/k_B = 259$ K, with k_B being the Boltzmann constant. The cross-interaction terms are given by the Lorentz−Berthelot rules.[5]

Since the smaller interaction site (H) is fully enclosed within the larger interaction site (Cl), there is a strong physical screening of the pair correlation between hydrogen interaction sites. As shown in Figure 3, the resulting behavior for $g_{HH}(r)$ (circle symbols) is similar to that of uncorrelated sites. This characteristic property of this model system is correctly reproduced by our solution. In this sense, it successfully eliminates the unphysical behavior exhibited by PISM-HNC approximation, in particular the negative region appearing in $g_{HH}(r)$ and $g_{HCL}(r)$ pair probability distribution functions and a peak predicted at about 0.4 Å in $g_{HH}(r)$. The reason probably lies in the fact that the different contributions to our correlation functions over those regions are appropriately cancelled when the optimal parametrization is reached.

In contrast, Lue and Blankschtein[27] showed that the aforementioned cancellation is not accomplished by adding a certain set of bridge diagrams to the PISM-HNC theory. In fact, this problem is also found in other known approximate theories based on diagrammatically proper integral equations. The chlorine−chlorine pair distribution function predicted by PISM-HNC is otherwise fairly similar to our solution, both being in good agreement with MD simulation. It is a consequence of the particular geometry exhibited by this fluid which makes the correlation between chlorine interaction sites quite affected by the internal structure of the molecules. As a result, the pair correlation function looks like the one obtained from atomic Lennard-Jones fluids where the thermodynamics and structure are rather insensitive to the specific approximation for the closure relationship.

B. Waterlike Triatomics. The second system analyzed in this paper is that of triatomic molecules characterized by the density $\rho = 0.03345$ Å^{-3}, screened density $\eta = \rho/1.0$,

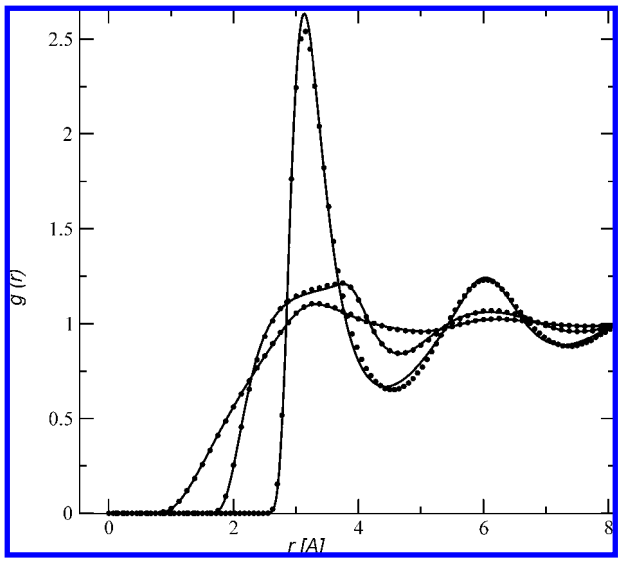

Figure 4. SPC model for water without charges: Site−site pair correlation functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ at density $\rho = 0.03345$ Å^{-3}, screened density $\eta = \rho/1.0$, and temperature $T = 298$ K. Solid line and circles correspond to this work and data simulation, respectively.

and temperature $T = 298$ K. The molecular geometry and Lennard-Jones parameters are that of the SPC model for water given elsewhere.[26,48] Initially, we examined the system without charges, $q_H = q_O = 0$, and later with charges.

These rigid models represent other interesting fluids for testing our approximate theory on nonpolar fluids since the internal structure of the molecule is completely different than the one analyzed previously. In fact, two intramolecular correlations or "bonds" ($L_{HO} = 1.0$ Å and $L_{HH} = 1.633$ Å) are strongly reflected in the behavior of the pair correlation functions. Also for this class of models as originally parametrized, the excluded volume or van der Waals intermolecular interactions (Lennard-Jones terms) act on only one species, namely the oxygen sites ($\sigma_{OO} = 3.1655$ Å and $\epsilon_{OO}/k_B = 78.13$ K).

As has been done before by comparing the structure of this nonpolar fluid with that of more realistic waterlike models with site charges, we are also able to explicitly show how charges, and specially hydrogen-bonds, affect the behavior of correlation functions between hydrogen and oxygen atoms.[49]

We plot in Figure 4 the site−site pair correlation functions predicted for each of $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$. We find remarkable agreement between our approximate solution (solid line) and MD simulation (circle symbols) over the entire range of distances. We found that our prediction reproduces very well not only the location and height of the first peaks but also the phase of oscillation at intermediate distances.

Next, we reanalyzed the same fluid with the addition of charges, $q_H = +0.41e$ for hydrogen and $q_O = -2q_H$ for oxygen, where e is the elementary charge. Since there would be an attractive pole when a positive and negative charge overlap, we modified the SPC model by introducing a small van der Waals (r^{-12}) repulsive hydrogen-oxygen interaction $u_{OH}(r) = A_{OH}[(kcal\ A^{12})/mol]r^{-12}$ to make a repulsive overlap of hydrogen and oxygen sites of different waterlike molecules for small intersite distances. This is an old patch for these model types.[49] We used two different values for the parameter A_{OH}, namely $A_{OH} = 900$ (Lue and Blankschtein's proposal)[50,51] and $A_{OH} = 225.18$ (Pettitt and Rossky's proposal).[49] The latter was designed to do minimal damage to the thermodynamics and structure of the original model whereas the former was apparently chosen on other grounds.

As a preliminary test of our optimized approach for polar multicomponent fluids, we used the PISM integral equation ($\eta = \rho$) combined with the interpolating closure relationship 10 to compute the structure. As described elsewhere,[26,52] the long-range nature of correlation functions is then approximately handled by dividing the direct correlation functions $c^o_{\alpha\gamma}(r)$ into a short-range and a long-range part and may be handled by renormalization or with other explicit analytical expressions by the method of Ng.[53] Note that, otherwise, expression 4 is no longer useful with long-range interactions. Indeed, the screened density approach also has to be appropriately renormalized in order to lead divergent contributions coming from the integration of those Coulomb terms that appear in the expansion of the partition function. Such extension of the theory is not trivial, and we leave the approach for a future article.

We plotted in Figure 5 Lue and Blankschtein's model and in Figure 6 Pettitt and Rossky's model for the site−site pair correlation functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ predicted by RISM-HNC (dashed line), PISM-HNC (dashed-dotted line), this work (solid line), and MD simulation (symbols). As expected the plots clearly show a significant change between the structure obtained with charges versus the uncharged case. An essential feature of water which is well represented by our prediction is given by the first peak in $g_{OH}(r)$ corresponding to the presence of hydrogen-bonding. Another characteristic of water is the narrow first peak in $g_{OO}(r)$ corresponding to a small coordination number, which is found to be very well reproduced by all of the theoretical predictions analyzed in this article. On the other hand, they do less in describing the long-range structure.

The optimization process is performed, as described in ref 30, by minimizing the approximate analytic expression obtained for the molecular excess chemical potential μ_{ex} 5. The independent parameters satisfying the physical symmetry properties of both models are a^o_{11}, a^o_{12}, a^o_{22}, a^r_{11}, a^r_{12}, a^r_{21}, a^r_{22}, a^b_{11}, a^b_{12}, and a^b_{22}, in which the label number 1 corresponds to hydrogen and the label number 2 corresponds to either chlorine or oxygen, for the first or the second model, respectively. The values of the parameters, obtained from the direct application of this procedure to those models, are the following: $a^o_{HH} = 16.00$, $a^o_{HCl} = 4.85$, $a^o_{ClCl} = 0.015$, $a^r_{HH} = 15.00$, $a^r_{HCl} = 6.00$, $a^r_{ClH} = 6.00$, $a^r_{ClCl} = 0.70$, $a^b_{HH} = 15.00$, $a^b_{HCl} = 1.00$, and $a^b_{ClCl} = 0.70$ for the nonpolar chloride-like model, $a^o_{HH} = 0.00$, $a^o_{HO} = 0.00$, $a^o_{OO} = 0.95$, $a^r_{HH} = 0.00$, $a^r_{HO} = 0.00$, $a^r_{OH} = 0.00$, $a^r_{OO} = 0.00$, $a^b_{HH} = 0.00$, $a^b_{HO} = 0.00$, and $a^b_{OO} = 0.00$ for the SPC model without charges, $a^o_{HH} = 0.025$, $a^o_{HO} = 1.5$, $a^o_{OO} = 1.27$, $a^r_{HH} = 0.5$, $a^r_{HO} = 1.5$, $a^r_{OH} = 1.5$, $a^r_{OO} = 1.27$, $a^b_{HH} = 0.09$, $a^b_{HO} = 1.5$, and $a^b_{OO} = 1.27$ for the modified SPC model for waterlike molecules proposed by Lue and Blankschtein, and

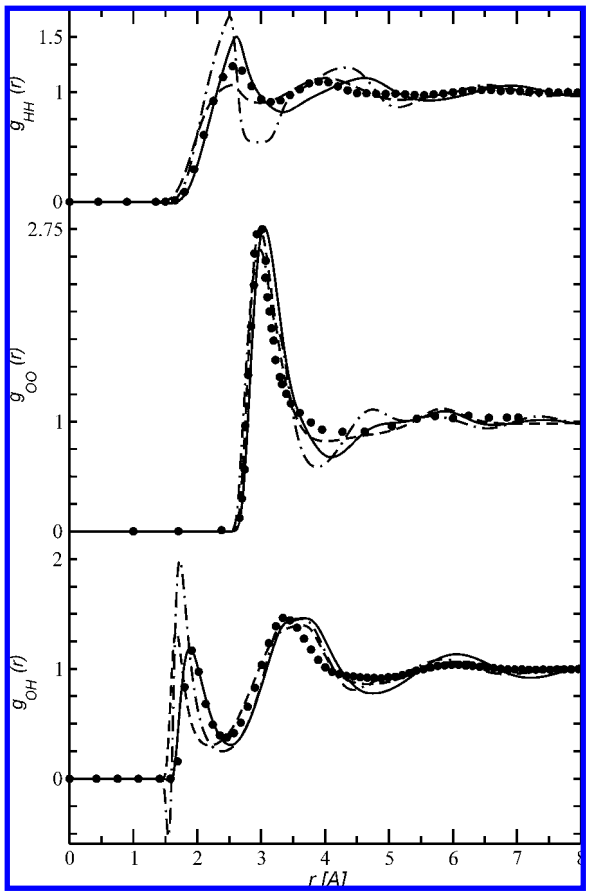

Figure 5. Modified SPC model for water proposed by Blue and Blankschtein: Site−site pair correlation functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ at density $\rho = 0.03345$ Å^{-3}, screened density $\eta = \rho$, and temperature $T = 298$ K. Dashed, dashed-dotted, and solid lines correspond to RISM-HNC, PISM-HNC, and this work, respectively. Circles represent MD simulation.

$a^o_{HH} = 0.07$, $a^o_{HO} = 1.58$, $a^o_{OO} = 0.58$, $a^r_{HH} = 1.01$, $a^r_{HO} = 1.55$, $a^r_{OH} = 1.47$, $a^r_{OO} = 1.15$, $a^b_{HH} = 0.04$, $a^b_{HO} = 1.53$, and $a^b_{OO} = 1.32$ for the modified SPC model for waterlike molecules proposed by Pettitt and Rossky.

Finally, we have tabulated the numerical results for the excess internal energy[40] U^{ex}/Nk_BT predicted by this work, PISM-HNC approximation, and MD simulation. This provides a test of these approximate theories and their solutions on relevant thermodynamic properties of nonpolar molecular fluids. As shown in Table 4, we found that our predictions are in better agreement with simulation data than the older theories.

V. Conclusions

We have introduced a thermodynamically consistent theory which has been successfully tested on representative models of heteronuclear molecular fluids, including water. It is based on an extension of that recently developed for homonuclear fluids. The closure approximation is obtained by analogy to one-component fluids. The approximate theory is completed by coupling the closure approximation with a variant of the diagrammatically proper integral equation introduced recently by this laboratory.

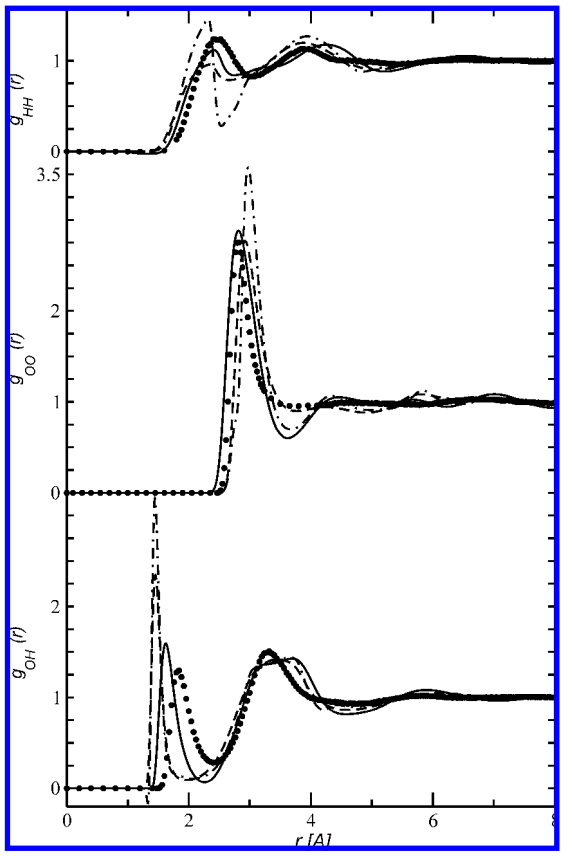

Figure 6. Modified SPC model for water proposed by Pettitt and Rossky: Site−site pair correlation functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ at density $\rho = 0.03345$ Å^{-3}, screened density $\eta = \rho$, and temperature $T = 298$ K. Dashed, dashed-dotted, and solid lines correspond to RISM-HNC, PISM-HNC, and this work, respectively. Circles represent MD simulation.

Table 4. Comparison between MD Simulation and Theoretical Predictions for the Excess Internal Energy U^{ex}/Nk_BT [a]

	PISM-HNC		this work		MD
molecule	U^{ex}/Nk_BT	$\beta\mu_{ex}$	U^{ex}/Nk_BT	$\beta\mu_{ex}$ [b]	U^{ex}/Nk_BT
HCl	−82.66	−4, 61	−87.43	−5.0	−88.22
H_2O	−6.71	11.87	−6.70	10.05	−6.45

[a] The excess chemical potential $\beta\mu_{ex}$ predicted by several approximate theories is also included in this table. [b] Results coming from the numerical evaluation of the Morita and Hiroike expression.

The optimal values of the parameters involved in this theory were obtained, as described in ref 30, by minimizing an approximate analytic expression obtained for the molecular excess chemical potential μ_{ex}. The systematic evaluation of correlation functions required by this optimization procedure required better algorithms to solve these integral equations. We introduced a high-performance algorithm to compute iteratively correlation functions, obtaining a significant reduction in the computational cost. Not surprisingly, it is found to be faster and more efficient than the direct Picard iteration.

This numerical method is a modification to the algorithm developed initially by Gillan for atomic fluids.[43] It consists basically of expanding the indirect correlation function on the sine function basis set, a natural expansion within a

scheme that already involves a systematic evaluation of sine Fourier transforms. A fast and robust Newton-GMRES algorithm based on Krylov's approach,[47] which approximates the solution for the weight of the projections in each iteration, is initially utilized to solve the set of nonlinear integral equations for the first M sine Fourier components (primary contribution), keeping the remaining components fixed. This is followed by one direct Picard iteration to refine the higher sine Fourier components, keeping those first components fixed. This process is repeated to get convergence.

The remarkable efficiency of this scheme is primarily due to the fact that less than 1% of the sine function basis set was sufficient to provide a good representation of the indirect correlation functions at a coarse level. The Newton−Raphson poor global convergence was notably improved by implementing a nested algorithm, which approximately solves the corresponding nonlinear equation systems using the previous approach on a sequence of meshes, ending with a solution at the target, finest mesh.[47] It was found that this method provides a fast and efficient process by which a sufficiently good initial guess for the target grid is reached at a very low computational cost.

This computational scheme was utilized to analyze the accuracy and efficiency of the theory in predicting structural and thermodynamic properties on two geometrically different polyatomic models. The first fluid is that of nonpolar diatomic molecules of hydrogen chloride (HCL) in which the smaller interaction site (H) is fully enclosed by the larger interaction site (Cl), screening the pair correlation between hydrogen interaction sites. We showed that this optimized theory is capable of successfully describing this phenomenon, eliminating the unphysical behavior exhibited by other known approximate theories based on diagrammatically proper integral equations, including PISM-HNC approximation, in particular the negative region appearing in $g_{HH}(r)$ and $g_{HCL}(r)$ pair correlation functions and a peak predicted at about 0.4 Å in $g_{HH}(r)$.

The second system is the SPC model family for water. It was initially examined without charges since it represents another challenging model for testing our approximate theory on nonpolar fluids. Certainly the geometry and internal structure of this molecule is completely different than the one analyzed previously. Further, the structure of this nonpolar fluid was afterwards compared with that of waterlike explicitly showing how charges, and specially hydrogen bonds, affect the behavior of the correlation functions between hydrogen and oxygen atoms. We found that our predictions for this geometrically more intricate model are in excellent agreement with MD simulation.

As a preliminary test of our optimized approach for polar multicomponent fluids, we used PISM integral equation combined with our interpolating closure relationship to describe the structure of waterlike molecular fluid, instead of the aforementioned screened density integral equation, since the latter is no longer valid in the present fashion for long-range interactions. We compared different approximate integral equation theories against simulation, finding that our prediction describes quite well essential features of water such as the first peak in $g_{HO}(r)$ and $g_{OO}(r)$ which represent the hydrogen-bond and the coordination number, respectively. Certainly a notable improvement in the prediction of the PISM integral equation is obtained when properly combined with the interpolating closure approximation proposed in the present article, rather than with HNC or PY approximations. Beyond predicting the first shell of water successfully, this approximate theory fails, as many other approximate theories, in correctly describing the asymptotic behavior of the corresponding correlation functions, which plays a crucial role in the prediction of a reasonable dielectric theory.[80,55] In fact, most of the approximate integral equation theories are not consistent with the zeroth and second moment conditions and yield trivial and incorrect predictions for the static dielectric constant.[52,56,57]

Further work involves the renormalization of the screened density integral equation. As already tested on nonpolar fluids, this variant of the PISM integral equation has been shown to provide an accurate description of structure and thermodynamic properties when it is combined with the interpolating closure approximation. Thus, it represents a promising powerful tool to correctly capture structural, dielectric, and thermodynamic properties of polar multicomponent molecular fluids. It may also provide a proper framework of self-consistency to describe the underlying physics of gas−liquid-phase transition of molecular fluids where most of the approximate integral equations have been shown to predict the incorrect critical temperature as well as isothermal compressibility and constant-volume heat capacity.[58,59]

Another direction we will consider in the future derived from this paper is related to the development of a proper interpolating closure approximation for 3D integral equation theories. It is well-known that many intricate biological systems require more detailed information on the structure of molecular fluids than the one that can be reasonably obtained from using 1D approximate theories.[1] On such 3D grids, a new challenge will surely be the design of a new high performance algorithm based on the ideas proposed in this article. An important future aim will be computing 3D correlation functions as accurately as possible at low computational complexity.

Acknowledgment. M.M. and B.M.P. thank the NIH and the Robert A. Welch foundation for partial support. M.M. is also grateful for financial support from the World Laboratory Center for Pan-American Collaboration in Science and Technology. The work of C.T.K. was supported by NSF grant #DMS-0707220. Drs. Kippi Dyer, John Perkyns, Ka-yiu Wong, and Gillian Lynch are acknowledged for many helpful conversations.

Appendix

The computational scheme described in section III can be summarized as follows:

Algorithm 1: nested_it ($\mathcal{T}^{guess}$, $\Delta\xi_o$, j_{max}, L, m, η_o, η M, IE, CL, $\mathcal{T}^{output}$, $\mathcal{C}^{output}$)

-Define the coarsest grid: $j = 0$; $\Delta\xi = \Delta\xi_o$; $N = L/\Delta\xi + 1$;

- *Call* the numerical_solver ($\mathcal{T}^{guess}$, $\Delta\xi$, N, m, η_o, M, IE, CL, $\mathcal{T}^{output}$, $\mathcal{C}^{output}$) to obtain the approximate solution for

the indirect and direct correlation functions $\mathcal{T}^{output}$ and $\mathcal{C}^{output}$ over the coarsest grid, respectively;

while ($j < j_{max}$) **do**

-Define the next level $j = j + 1$;

-Increase gradually the precision η_j in such a way tha $\eta_{jmax} = \eta$;

-Define the input (initial guess) over the finer grid using the intergrid interpolation

$\mathcal{T}^{input}_{\mu\nu,2p} = \mathcal{T}^{output}_{\mu\nu,p}$ and $\mathcal{T}^{input}_{\mu\nu,2p+1} = 1/2(\mathcal{T}^{output}_{\mu\nu,p+1} + \mathcal{T}^{output}_{\mu\nu,p})$ for $p = 0, .., (N - 1)$ and $\mu, \nu = 1, .., 2m$;

-Define the new grid: $\Delta\xi = \Delta\xi/2$; $N = L/\Delta\xi + 1$;

-*Call* the numerical_solver ($\mathcal{T}^{input}$, $\Delta\xi$, N, m, η_j, M, *IE*, *CL*, $\mathcal{T}^{output}$,$\mathcal{C}^{output}$) to obtain the approximate solution for the indirect and direct correlation functions $\mathcal{T}^{output}$ and $\mathcal{C}^{output}$ over the finer grid, respectively;

end while

Hereafter $\mu, \nu = 1, .., 2m$; $p = 1, .., N - 1$; $n = 1, .., M$; $k = 1 + M, .., N - 1$ when required;

Algorithm 2: numerical_solver ($\mathcal{T}^{input}$, $\Delta\xi$, N, m, η, M, *IE*, *CL*, $\mathcal{T}^{output}$,$\mathcal{C}^{output}$)

-Use $\mathcal{T}^{input}_{\mu\nu,p}$ to obtain a new estimate for $\mathcal{T}^{output}_{\mu\nu,p}$ and $\mathcal{C}^{output}_{\mu\nu,p}$ from the direct evaluation of eqs 16, 13, 11, and 12, as performed by one Picard iteration;

-Define the parameter $a = \sqrt{\sum_{\mu,\nu,p}[\mathcal{T}^{output}_{\mu\nu,p} - \mathcal{T}^{input}_{\mu\nu,p}]^2}$;

while ($a > \sqrt{N-1}\eta$) **do**

$\mathcal{T}^{input}_{\mu\nu,p} = \mathcal{T}^{output}_{\mu\nu,p}$

-Use $\mathcal{T}^{output}_{\mu\nu,p}$ to compute the sine Fourier components $\hat{\mathcal{T}}_{\mu\nu,n}$ using the FFT 13;

-Use $\hat{\mathcal{T}}_{\mu\nu,n}$ and $\mathcal{T}^{output}_{\mu\nu,p}$ to compute the coarse and fine part as shown by expression 12 and keep the fine part contribution fixed;

-Use $\mathcal{T}^{output}_{\mu\nu,p}$ to obtain the direct correlation function $\mathcal{C}_{\mu\nu,p}$ from eq 16 and use it to obtain $\hat{\mathcal{C}}_{\mu\nu,p}$ using the FFT 13;

-Evaluate the Jacobian 17 at $\hat{\mathcal{T}}_{\mu\nu,n}$ and the residual 14 at $\hat{\mathcal{C}}_{\mu\nu,p}$ as described above and then, use them to solve eq 19 to estimate changes on the downhill direction over the first M sine Fourier components $\Delta\hat{\mathcal{T}}_{\mu\nu,n}$;

-Define $\delta = a/10$

while ($\sqrt{\sum_{\mu,\nu,n}[\Delta\hat{\mathcal{T}}_{\mu\nu,n}]^2} > \sqrt{M}\delta$) **do**

- Calculate the new first M sine Fourier components $\hat{\mathcal{T}}^{new}_{\mu\nu,n} = \Delta\hat{\mathcal{T}}_{\mu\nu,n} - \hat{\mathcal{T}}_{\mu\nu,n}$;

-Use $\hat{\mathcal{T}}^{new}_{\mu\nu,n}$ to compute the coarse part as shown by expression 12 and then get $\mathcal{T}^{output}_{\mu\nu,p}$ by keeping the fine part contribution fixed;

- Use $\mathcal{T}^{output}_{\mu\nu,p}$ to obtain the direct correlation function $\mathcal{C}_{\mu\nu,p}$ from eq 16 and use it to obtain $\hat{\mathcal{C}}^{new}_{\mu\nu,p}$ using the FFT 13;

- Evaluate the Jacobian 17 at $\hat{\mathcal{T}}^{new}_{\mu\nu,n}$ and the residual 14 at $\hat{\mathcal{C}}^{new}_{\mu\nu,p}$ as described above and then, use them to solve eq 19 to estimate new changes on the new downhill direction over the first M sine Fourier components $\Delta\hat{\mathcal{T}}^{new}_{\alpha\gamma,n}$;

$\hat{\mathcal{T}}_{\mu\nu,n} = \hat{\mathcal{T}}^{new}_{\alpha\gamma,n}$; $\Delta\hat{\mathcal{T}}_{\mu\nu,n} = \Delta\hat{\mathcal{T}}^{new}_{\alpha\gamma,n}$

end while

-Use $\hat{\mathcal{T}}_{\mu\nu,n}$ to compute the coarse part as shown by expression (12) and then get $\mathcal{T}^{inter}_{\mu\nu,p}$ by keeping the fine part contribution fixed;

-Use $\mathcal{T}^{inter}_{\mu\nu,p}$ to obtain a new estimate for $\mathcal{T}^{output}_{\mu\nu,p}$ and $\mathcal{C}^{output}_{\mu\nu,p}$ and the higher sine Fourier components $\hat{\mathcal{T}}_{\mu\nu,k}$ from the direct evaluation of eqs 16, 13, 11, and 12, as performed by one Picard iteration;

-Use $\hat{\mathcal{T}}_{\mu\nu,k}$ to get the new fine part contribution as shown by eq 12 by keeping the coarse part contribution fixed;

$a = \sqrt{\sum_{\mu,\nu,p}[\mathcal{T}^{output}_{\mu\nu,p} - \mathcal{T}^{input}_{\mu\nu,p}]^2}$

end while

References

(1) Hirata, F. *Molecular Theory of solvation*; Kluwer Academic: Amsterdam, 2004.

(2) Cramer, C. J.; Truhlar, D. G. Continuum solvation models: classical and quantum mechanical implementations. In *Reviews in computational chemistry;* Lipkowitz, K. B., Ed.; VCH: New York, 1995.

(3) Fantoni, R.; Pastore, G. *J. Chem. Phys.* **2003**, *119*, 3810.

(4) Hansen, J. P.; McDonald, I. R. *Theory of simple Liquids*; Academic Press Inc.: London, 1986.

(5) Lee, L. L. *Molecular Thermodynamics of Nonideal Fluids*; Butterworths: Boston, 1988.

(6) Percus, J. K.; Yevick, G. J. *Phys. Rev.* **1958**, *110*, 1.

(7) Stell, G. *Physica* **1963**, *29*, 517.

(8) Morita, T. *Prog. Theor. Phys.* **1960**, *23*, 829.

(9) Verlet, L. *Nuovo Cimento* **1960**, *18*, 77.

(10) Chandler, D. In *The Liquid State of Matter: Fluids, Simple and Complex;* Montroll, E. W., Lebowitz, J. L., Eds.; North Holland Pub. Co.: Amsterdam, 1982; p 275.

(11) Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1972**, *57*, 1930.

(12) Ladanyi, B. M.; Chandler, D. *J. Chem. Phys.* **1975**, *62*, 4308.

(13) Chandler, D.; Pratt, L. *J. Chem. Phys.* **1976**, *65*, 2925.

(14) Chandler, D. *Mol. Phys.* **1976**, *31*, 1213.

(15) Chandler, D. *J. Chem. Phys.* **1977**, *67*, 1113.

(16) Hirata, F.; Rossky, P. J.; Pettitt, B. M. *J. Chem. Phys.* **1983**, *78*, 4133.

(17) Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* **1983**, *78*, 7296 and references therein.

(18) Jhonson, E.; Hazoume, R. P. *J. Chem. Phys.* **1979**, *70*, 1599.

(19) Monson, P. *Mol. Phys.* **1982**, *47*, 435.

(20) Sullivan, D. E.; Gray, C. G. *Mol. Phys.* **1981**, *42*, 443.

(21) Cummings, P. T.; Stell, G. *Mol. Phys.* **1981**, *44*, 529.

(22) Chandler, D.; Silbey, R.; Ladanyi, B. *Mol. Phys.* **1982**, *46*, 1335.

(23) Chandler, D.; Joslin, C. G.; Deutch, J. M. *Mol. Phys.* **1982**, *47*, 871.

(24) Rossky, P. J.; Chiles, R. A. *Mol. Phys.* **1984**, *51*, 661.

(25) Lupkowski, M.; Monson, P. A. *J. Chem. Phys.* **1987**, *87*, 3618.

(26) Lue, L.; Blankschtein, D. *J. Chem. Phys.* **1995**, *102*, 5427.

(27) Lue, L.; Blankschtein, D. *J. Chem. Phys.* **1995**, *102*, 4203.

(28) Dyer, K. M.; Perkyns, J. S.; Pettitt, B. M. *J. Chem. Phys.* **2005**, *122*, 236101.

(29) Dyer, K. M.; Perkyns, J. S.; Pettitt, B. M. *J. Chem. Phys.* **2005**, *123*, 204512.

(30) Marucho, M.; Pettitt, B. M. *J. Chem. Phys.* **2007**, *78*, 7296.

(31) Lebowitz, J. L.; Percus, J. K. *Phys. Rev.* **1961**, *122*, 1675.

(32) Stell, G. *Mol. Phys.* **1969**, *16*, 209.

(33) Kinoshita, M.; Harada, M. *Mol. Phys.* **1991**, *74*, 443.

(34) Kinoshita, M.; Harada, M. *Mol. Phys.* **1993**, *79*, 145.

(35) Kinoshita, M.; Harada, M. *Mol. Phys.* **1994**, *81*, 1473.

(36) Halminton, T. P.; Pulay, P. *J. Chem. Phys.* **1986**, *84*, 5728.

(37) Kovalenko, A.; Ten-no, S.; Firata, F. *J. Comput. Chem.* **1999**, *20*, 928.

(38) Booth, M. J.; Schlijper, A. G.; Scales, L. E.; Haymet, D. J. *Comput. Phys. Commun.* **1999**, *119*, 122.

(39) Vatamann, J.; Cann, N. M. *J. Chem. Phys.* **2004**, *121*, 6922.

(40) Lue, L.; Blankschtein, D. *J. Chem. Phys.* **1994**, *100*, 3002.

(41) Ryzhik, G.; Jeffrey, A. *Table of Integrals, Series, and Products*, 5th ed.; Academic Press, Inc.: London, 1986.

(42) Abramowitz, M.; Stegun, I. A. *Handbook of Mathematical functions*; Dover: New York, 1972.

(43) Gillan, M. J. *Mol. Phys.* **1979**, *38*, 79.

(44) Labik, S.; Malijevsky, A.; Vonka, P. *Mol. Phys.* **1985**, *56*, 709.

(45) Ichiye, T.; Haymet, D. J. *J. Chem. Phys.* **1988**, *89*, 4315.

(46) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys.* **1985**, *54*, 1712.

(47) Kelley, C. T.; Pettitt, B. M. *J. Comput. Phys.* **2004**, *197*, 491.

(48) Glattli, A.; Daura, X.; Van Gunsteren, W. F. *J. Chem. Phys.* **2002**, *116*, 9811.

(49) Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* **1982**, *77*, 1452.

(50) Lue, L.; Blankschtein, D. *J. Chem. Phys.* **1992**, *96*, 8582.

(51) Redy, G.; Lawrence, C. P.; Skinner, J. L.; Yethiraj, A. *J. Chem. Phys.* **2003**, *119*, 13012.

(52) Chandler, D.; Joslin, G. S.; Deutch, J. M. *Mol. Phys.* **1982**, *47*, 871.

(53) Ng, K.-C. *J. Chem. Phys.* **1974**, 61, 2680.

(54) Chandler, D. *J. Chem. Phys.* **1977**, *67*, 1113.

(55) Hoye, J. S.; Stell, G. *J. Chem. Phys.* **1976**, *65*, 18.

(56) Perkyns, J.; Pettitt, B. M. *Chem. Phys. Lett.* **1992**, *190*, 626.

(57) Perkyns, J.; Pettitt, B. M. *J. Chem. Phys.* **1992**, *97*, 7656.

(58) Sarkisov, G.; Lomba, E. *J. Chem. Phys.* **2005**, *122*, 214504.

(59) Peplow, A. T.; Beardmore, R. E.; Bresme, F. *Phys. Rev. E* **2006**, *74*, 1539.

CT700202H