A reference-modified density functional theory: an application to solvation free-energy calculations for a Lennard-Jones solution

Tomonari Sumi,[1,a)] Yutaka Maruyama,[2] Ayori Mitsutake,[3,4] and Kenichiro Koga[1]

[1] *Department of Chemistry, Faculty of Science, Okayama University, 3-1-1 Tsushima-Naka, Kita-ku, Okayama 700-8530, Japan*

[2] *Co-Design Team, Exascale Computing Project, RIKEN Advanced Institute for Computational Science, 7-1-26, Minatojima-minami-machi, Kobe 650-0047, Japan*

[3] *Department of Physics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223–8522, Japan*

[4] *JST, PREST, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223–8522, Japan*

In the conventional classical density functional theory (DFT) for simple fluids, an ideal gas is usually chosen as the reference system because there is a one-to-one correspondence between the external field and the density distribution function, and the exact intrinsic free-energy functional is available for the ideal gas. In this case, the second-order density functional Taylor series expansion of the excess intrinsic free-energy functional provides the hypernetted-chain (HNC) approximation. Recently, it has been shown that the HNC approximation significantly overestimates the solvation free energy (SFE) for an infinitely-dilute Lennard-Jones (LJ) solution, especially when the solute particles are several times larger than the solvent particles [T. Miyata and J. Thapa, Chem. Phys. Lett. **604**, 122 (2014)]. In the present study, we propose a reference-modified density functional theory (RMDFT) as a systematic approach to improve the SFE functional as well as the pair distribution functions. The second-order density functional Taylor series expansion for the excess part of the intrinsic free-energy functional, in which a hard-sphere fluid is introduced as the reference system instead of an ideal gas, is applied to the LJ pure and infinitely-dilute solution systems, and is proved to remarkably improve the drawbacks of the HNC approximation. Furthermore, the third-order density functional expansion approximation, in which a factorization approximation is applied to the triplet direct correlation function, is examined for the LJ systems. We also show that the third-order contribution can yield further refinements for both the pair distribution function and the excess chemical potential for the pure LJ liquids.

I. INTRODUCTION

Solvation free energy (SFE), which is defined as the change in free energy accompanied with the transfer of a solute molecule from its dilute gas to a solution system with the same number density as in the dilute gas, is one of the most important thermodynamic quantities in solution chemistry[1] because the SFE is directly related to the solubility of the dilute gas. Furthermore, the thermodynamic stability of large complex solute molecules, such as proteins in various conformations, is determined by a difference in the SFE resulting from conformational changes. Thus, an accurate theoretical prediction of the SFE is one of the most important goals in computational physical chemistry.

a) Electronic mail: sumi@okayama-u.ac.jp.

The standard approach for calculating the SFE using molecular simulations is based on either the free-energy perturbation method or the thermodynamic integration method.[2,3] These calculations using molecular simulations are computationally exact [4-7]; however, they are time consuming because molecular simulations should be performed for a number of intermediate states in the process of growing a solute molecule in the solution. To avoid having to run molecular simulations for the intermediate states, the energy representation (ER) method has been proposed.[8-10] However, since the ER method also requires two molecular simulation runs for the pure solvent and solution system, it is time consuming to apply this method to SFE calculations for large and complex systems such as biomolecules.

Alternatively, SFE calculations can be done by applying statistical mechanical approaches based on integral equation theories [11-19] and density functional theories[20,21] to avoid extensive numerical simulations. One of the most popular integral equation theories for molecular liquids is the extended reference-interaction-site model (RISM) theory,[13,22] where the hypernetted-chain (HNC)-type equation is employed as the closure relation. However, the use of the HNC equation as the closure for the RISM equation is not theoretically justified by statistical mechanics, because the HNC equation was originally derived as the closure relation not for the RISM (or site-site Ornstein-Zernike (OZ)) equation of polyatomic molecular liquids but for the OZ equation of simple liquids consisting of spherical particles.[23]

It has been pointed out that the SFE determined using the RISM equation with the HNC-type closure artificially depends on the number of sites on the solute molecules. [24,25] The dummy site problem on the SFE pointed out by Ten-no, [24] where the SFE is artificially increased by an addition of dummy site on solute molecule, would be separated from the artificially large increase in the SFE with the increase in either the number of sites or molecular size of solute. The dummy site problem is basically attributed to the theoretical drawback of the 1-D RISM equation [13,22] for solute molecules, thus it can be resolved by the three-dimensional treatment of the solute molecule based on the three-dimensional reference-interaction-site model (3D-RISM) integral equation. [16,26] In the similar point of view, the significant large overestimation in the SFE for large solute molecules using the RISM approach had been considered as one of the theoretical problems arising from the extension of the HNC closure relation for simple liquids to the interaction-site model for polyatomic molecular liquids. However, very recently, Miyata and Thapa showed that the SFE of a Lennard-Jones (LJ) solute in a LJ solvent, which was calculated using the HNC or Kovalenko–Hirata (KH) closure, monotonically increases as the size of solute increases, although the SFE determined using molecular dynamics (MD) simulations decreases as the size of solute increases.[27] Miyata's study suggests that the qualitatively incorrect solute-size dependence given by the RISM theory should be attributed to a theoretical drawback of the HNC-type approximation rather than to the simple extension of the HNC approximation to the interaction-site model of polyatomic molecular liquids. In fact, we demonstrated that a SFE functional derived properly from a second-

order density functional Taylor series expansion, namely, a HNC-type approximation that specializes in polyatomic molecular liquids, also provides artificially large solute-size dependence on the SFE for methane, propane, and isobutane in water. [28,29] Thus, it is expected that the theoretical prediction of the SFE based on integral equations is significantly improved, if the HNC approximation is replaced by a post-HNC theory.

The HNC equation can be derived via a diagrammatic analysis of the pair distribution function, if one neglects a set of closely connected diagrams, called the bridge functions.[23] Rosenfeld and Ashcroft pointed out that the bridge functions do not depend on the details of the interaction potential and, thus, should have an approximately universal function form.[30] It is also well known that the HNC equation can be derived via density functional theory (DFT).[23,31] Here, we consider an inhomogeneous simple liquid under an arbitrary external field $U(\mathbf{r})$. The grand potential under the external field is given by $\Omega[U] = -(1/\beta)\ln \Xi[U]$, where $\Xi[U]$ is the grand canonical partition function and $1/\beta$ is the Boltzmann's constant multiplied by the temperature, $k_B T$. The relation between the grand potential $\Omega[U]$ and the intrinsic free-energy density functional $F[n]$ is given by[32]

$$\Omega[U] = F[n] + \int d\mathbf{r}_1 n(\mathbf{r}_1|U)\left[U(\mathbf{r}_1) - \mu\right], \tag{1}$$

where μ is the chemical potential and $n(\mathbf{r}|U)$ is the density distribution function under the external field. Here, $n(\mathbf{r}|U)$ is also given by the functional differentiation of $\Omega[U]$:

$$n(\mathbf{r}|U) = \left.\frac{\delta\Omega[U]}{\delta\left[U(\mathbf{r}) - \mu\right]}\right|_{T,V}. \tag{2}$$

The intrinsic free-energy functional $F[n]$ should satisfy the Euler-Lagrange equation:

$$\left.\frac{\delta F[n]}{\delta n(\mathbf{r}|U)}\right|_{T,V} = \mu - U(\mathbf{r}). \tag{3}$$

Note that $U(\mathbf{r})$ in the Euler-Lagrange equation of Eq. (3) should be a density functional, i.e., $U(\mathbf{r}|n)$. In the conventional classical DFT for simple fluids, an ideal gas (IG) is usually introduced as the reference system. Thus, we assume that the density distribution function for the system of interest can be reproduced using an effective external field on the reference IG, $U_{IG}(\mathbf{r})$:

$$n_{IG}(\mathbf{r}|U_{IG}) = n(\mathbf{r}|U), \tag{4}$$

where

$$n_{IG}\left(\mathbf{r}\middle|U_{IG}\right)=n_0\exp\left[-\beta U_{IG}\left(\mathbf{r}\right)\right]. \tag{5}$$

In Eq. (5), $n_0=\exp\left(\beta\mu_{IG}\right)\Lambda^{-3}$, where μ_{IG} is the chemical potential for an ideal gas and Λ is the de Broglie thermal wavelength. Note that Eq. (4) can be regarded as the definition of $U_{IG}\left(\mathbf{r}\right)$. The intrinsic free energy functional for the reference IG system, $F_{IG}\left[n_{IG}\right]$, should also satisfy the Euler-Lagrange equation given by

$$\left.\frac{\delta F_{IG}\left[n_{IG}\right]}{\delta n_{IG}\left(\mathbf{r}\middle|U_{IG}\right)}\right|_{T,V}=\mu_{IG}-U_{IG}\left(\mathbf{r}\right). \tag{6}$$

It is obvious from Eq. (5) that $U_{IG}\left(\mathbf{r}\right)$ is also the density functional, i.e., $U_{IG}\left(\mathbf{r}\middle|n_{IG}\right)$. From Eqs. (5) and (6), we obtain the exact expression for $F_{IG}\left[n_{IG}\right]$:

$$\begin{aligned}F_{IG}\left[n_{IG}\right]=F_{IG}\left[n_0\right]+\left(\mu_{IG}-\frac{1}{\beta}\right)\int d\mathbf{r}_1\left[n_{IG}\left(\mathbf{r}_1\middle|U_{IG}\right)-n_0\right]\\ -\int d\mathbf{r}_1 n_{IG}\left(\mathbf{r}_1\middle|U_{IG}\right)U_{IG}\left(\mathbf{r}_1\right)\end{aligned}, \tag{7}$$

where

$$F_{IG}\left[n_0\right]=\left(\mu_{IG}-\frac{1}{\beta}\right)N. \tag{8}$$

The excess intrinsic free energy functional is defined as the difference between the intrinsic free energy functional $F\left[n\right]$ and that for the reference IG system:

$$F^{ex}\left[n\right]=F\left[n\right]-F_{IG}\left[n\right], \tag{9}$$

where n_{IG} is given by Eq. (4).

Now, we apply the density functional Taylor series expansion to $F^{ex}\left[n\right]$ with respect to the bulk density n_0:

$$\begin{aligned}F^{ex}\left[n\right]=F^{ex}\left[n_0\right]+\mu^{ex}\int d\mathbf{r}_1\Delta n\left(\mathbf{r}_1\middle|U\right)\\ -\frac{1}{2\beta}\int d\mathbf{r}_1\,d\mathbf{r}_2 C^{(2)}\left(\left|\mathbf{r}_1-\mathbf{r}_2\right|\right)\Delta n\left(\mathbf{r}_1\middle|U\right)\Delta n\left(\mathbf{r}_2\middle|U\right)\\ -\frac{1}{6\beta}\int d\mathbf{r}_1\,d\mathbf{r}_2\,d\mathbf{r}_3 C^{(3)}\left(\left|\mathbf{r}_1-\mathbf{r}_2\right|,\left|\mathbf{r}_2-\mathbf{r}_3\right|,\left|\mathbf{r}_3-\mathbf{r}_1\right|\right)\Delta n\left(\mathbf{r}_1\middle|U\right)\Delta n\left(\mathbf{r}_2\middle|U\right)\Delta n\left(\mathbf{r}_3\middle|U\right)\\ +\cdots\end{aligned} \tag{10}$$

where $\Delta n(\mathbf{r}|U) = n(\mathbf{r}|U) - n_0$,

$$\mu^{ex} = \left.\frac{\delta F^{ex}[n]}{\delta n(\mathbf{r}|U)}\right|_{T,V,U=0}, \tag{11}$$

$$C^{(2)}(|\mathbf{r}-\mathbf{r}'|) = -\beta \left.\frac{\delta^2 F^{ex}[n]}{\delta n(\mathbf{r}|U)\delta n(\mathbf{r}'|U)}\right|_{T,V,U=0}, \tag{12}$$

and

$$C^{(3)}(|\mathbf{r}-\mathbf{r}'|,|\mathbf{r}'-\mathbf{r}''|,|\mathbf{r}''-\mathbf{r}|) = -\beta \left.\frac{\delta^3 F^{ex}[n]}{\delta n(\mathbf{r}|U)\delta n(\mathbf{r}'|U)\delta n(\mathbf{r}''|U)}\right|_{T,V,U=0}. \tag{13}$$

Equation (12) is the second-order direct correlation function that is related to the pair correlation function, $h^{(2)}(|\mathbf{r}|)$, via the OZ equation (defined by Eq. (37)[23]) and Eq. (13) is the triplet direct correlation function that is related to the triplet correlation function, $h^{(3)}(|\mathbf{r}|,|\mathbf{r}'|,|\mathbf{r}+\mathbf{r}'|)$, via the OZ3 equation [see APPENDIX A].[33] From Eqs. (3), (4), (6), (9), (10), and (11), we obtain the following effective external field for the reference IG system:

$$\begin{aligned} U_{IG}(\mathbf{r}) = U(\mathbf{r}) &- \frac{1}{\beta}\int d\mathbf{r}_1 C^{(2)}(|\mathbf{r}-\mathbf{r}_1|)\Delta n(\mathbf{r}_1|U) \\ &- \frac{1}{2\beta}\int d\mathbf{r}_1 d\mathbf{r}_2 C^{(3)}(|\mathbf{r}-\mathbf{r}_1|,|\mathbf{r}_1-\mathbf{r}_2|,|\mathbf{r}_2-\mathbf{r}|)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U) \\ &+ \cdots \end{aligned}. \tag{14}$$

If we neglect the terms that are higher-order than the $C^{(2)}\Delta n$ term in Eqs. (10) and (14), these equations are reduced to the HNC approximation in which the bridge function is neglected.[23] This simplification occurs because the $C^{(2)}\Delta n$ term in Eq. (14) can be replaced by $h^{(2)}(r) - C^{(2)}(r)$ from Percus' relation (Eq. (36)) and the OZ equation (Eq. (37)). As a result, Eqs. (4), (5), and (14) give the HNC closure. Therefore, the $C^{(3)}$ term, as well as the higher-order terms, can be regarded as a part of the bridge function. Note that two assumptions are introduced in the HNC approximation. One is the use of an ideal gas as the reference system in the DFT and the second is the second-order truncation of the density functional Taylor series expansion of the excess free energy functional, $F^{ex}[n] = F[n] - F_{IG}[n]$. However, convergence of the density functional Taylor series expansion is not sufficiently rapid to justify the second-order truncation, e.g., in the HNC approximation for dense hard-sphere fluids.[34,35]

The aim of the present study is to present a general formulation for systematically developing the SFE functional and the integral equations for pure and solution systems consisting of simple spherical particles and polyatomic molecules. In order to examine the validity of the theory, we applied it to LJ pure and solution systems and then compared the numerical results with those of molecular dynamics (MD) simulations.

The rest of the paper is organized as follows. In Section II, the reference-modified density functional theory (RMDFT) is introduced, and the theoretical details for deriving the SFE functional and integral equations for pair distribution functions are described. In Section III, the method for calculating pair distribution functions and SFE is presented. In Section IV, the pair distribution functions for the LJ pure and solution systems calculated using these theories are compared with the results provided by MD simulations. A comparison of the theoretical results of the SFE with MD simulation data is also presented and discussed. Finally, our conclusions are presented in Section V.

II. REFERENCE-MODIFIED DENSITY FUNCTIONAL THEORY

A. Free-energy density functional

In the conventional classical DFT for simple fluids, an ideal gas is normally introduced as the reference system because the exact intrinsic free-energy functional, $F_{IG}[n]$, (given by Eq. (7)) is available for the ideal gas and the one-to-one correspondence between the density distribution function, $n_{IG}\left(\mathbf{r}\middle|U_{IG}\right)$, and the external field, $U_{IG}(\mathbf{r})$, (Eq. (5)) are also available. However, the most crucial interaction in dense liquids is the excluded volume interaction that can be qualitatively modeled as a hard sphere interaction. Thus, an ideal gas is ineffective as the reference system for constructing a free-energy functional for dense classical liquids if a density-functional Taylor series expansion such as the HNC approximation is used, even though the approximation with the ideal gas can provide the exact solution for the low-density limit of classical liquids. In this study, we propose a reference-modified density functional theory (RMDFT) in which an effective reference system is chosen for the system of interest to construct the free-energy density functional and the integral equation for density distribution function under arbitrary external field. When classical simple liquids are of interest, a hard sphere (HS) fluid is taken to be the reference system, instead of an ideal gas, to improve the HNC approximation. [See Fig. 1].

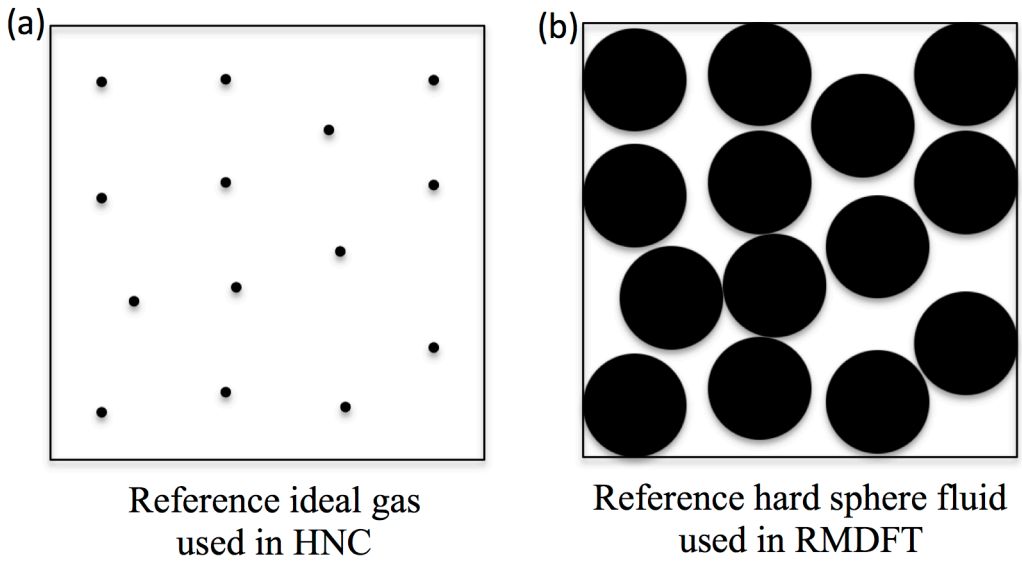


FIG. 1 Comparison of the reference systems used in classical DFT. (a) Ideal gas as the reference system employed in the HNC approximation. (b) Hard-sphere fluid as the reference system introduced in RMDFT. The hard-sphere fluid is a better reference system for dense classical liquids.

We can use an HS fluid, instead of ideal gas, as the reference system for the DFT because accurate free-energy density functional models are available for HS fluids.[35-38] Then, the excess part of the intrinsic free energy functional is, instead of Eq. (9), taken to be

$$\overline{F}^{ex}[n] = F[n] - F_{HS}[n], \tag{15}$$

Note that, in Eq. (15), we assume the following ansatz in which the density distribution function for the system of interest can be reproduced via that of the reference HS system by introducing an effective external field on the reference HS system, $U_{HS}(\mathbf{r})$:

$$n_{HS}(\mathbf{r}|U_{HS}) = n(\mathbf{r}|U), \tag{16}$$

where $U_{HS}(\mathbf{r})$ satisfies the following Euler-Lagrange equation for the reference HS system:

$$\left.\frac{\delta F_{HS}[n]}{\delta n(\mathbf{r}|U)}\right|_{T,V} = \mu_{HS} - U_{HS}(\mathbf{r}). \tag{17}$$

Here, Eq. (16) is used in Eq. (17). From Eqs. (3) and (17), we obtain

$$U_{HS}(\mathbf{r}) = U(\mathbf{r}) + \left.\frac{\delta \overline{F}^{ex}[n]}{\delta n(\mathbf{r}|U)}\right|_{T,V} - \overline{\mu}^{ex}, \tag{18}$$

where

$$\overline{\mu}^{ex} = \mu - \mu_{HS} \,. \tag{19}$$

From Eqs. (6) and (17), we obtain

$$U_{IG}(\mathbf{r}) = U_{HS}(\mathbf{r}) + \left.\frac{\delta F_{HS}^{ex}[n]}{\delta n(\mathbf{r}|U)}\right|_{T,V} - \mu_{HS}^{ex} \,, \tag{20}$$

where

$$F_{HS}^{ex}[n] = F_{HS}[n] - F_{IG}[n] \,, \tag{21}$$

and

$$\mu_{HS}^{ex} = \mu_{HS} - \mu_{IG} \,. \tag{22}$$

In Eq. (20), we also assume that

$$n_{IG}(\mathbf{r}|U_{IG}) = n_{HS}(\mathbf{r}|U_{HS}) \,. \tag{23}$$

Here, in the same manner as in Eq. (10), we apply a density functional Taylor series expansion to the excess part of the intrinsic free energy functional $\overline{F}^{ex}[n]$ redefined by Eq. (15):

$$\begin{aligned}\overline{F}^{ex}[n] = {} & \overline{F}^{ex}[n_0] + \overline{\mu}^{ex}\int d\mathbf{r}_1 \Delta n(\mathbf{r}_1|U) \\ & - \frac{1}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(2)}(|\mathbf{r}_1 - \mathbf{r}_2|)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U) \\ & - \frac{1}{6\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 C_{ex}^{(3)}(|\mathbf{r}_1 - \mathbf{r}_2|,|\mathbf{r}_2 - \mathbf{r}_3|,|\mathbf{r}_3 - \mathbf{r}_1|)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U)\Delta n(\mathbf{r}_3|U) \\ & + \cdots \end{aligned} \,, \tag{24}$$

where

$$\overline{\mu}^{ex} = \left.\frac{\delta \overline{F}^{ex}[n]}{\delta n(\mathbf{r}|U)}\right|_{T,V,U=0} \,, \tag{25}$$

$$C_{ex}^{(2)}(|\mathbf{r} - \mathbf{r}'|) = -\beta \left.\frac{\delta^2 \overline{F}^{ex}[n]}{\delta n(\mathbf{r}|U)\delta n(\mathbf{r}'|U)}\right|_{T,V,U=0} = C^{(2)}(|\mathbf{r} - \mathbf{r}'|) - C_{HS}^{(2)}(|\mathbf{r} - \mathbf{r}'|), \tag{26}$$

and

$$C_{ex}^{(3)}\left(|\mathbf{r}-\mathbf{r}'|,|\mathbf{r}'-\mathbf{r}''|,|\mathbf{r}''-\mathbf{r}|\right)=-\beta\left.\frac{\delta^3\overline{F}^{ex}[n]}{\delta n(\mathbf{r}|U)\delta n(\mathbf{r}'|U)\delta n(\mathbf{r}''|U)}\right|_{T,V,U=0}$$
$$=C^{(3)}\left(|\mathbf{r}-\mathbf{r}'|,|\mathbf{r}'-\mathbf{r}''|,|\mathbf{r}''-\mathbf{r}|\right)-C_{HS}^{(3)}\left(|\mathbf{r}-\mathbf{r}'|,|\mathbf{r}'-\mathbf{r}''|,|\mathbf{r}''-\mathbf{r}|\right). \quad (27)$$

The effective external field for the reference HS system given by Eq. (18) can be rewritten using Eq. (24):

$$U_{HS}(\mathbf{r})=U(\mathbf{r})-\frac{1}{\beta}\int d\mathbf{r}_1 C_{ex}^{(2)}\left(|\mathbf{r}-\mathbf{r}_1|\right)\Delta n(\mathbf{r}_1|U)$$
$$-\frac{1}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(3)}\left(|\mathbf{r}-\mathbf{r}_1|,|\mathbf{r}_1-\mathbf{r}_2|,|\mathbf{r}_2-\mathbf{r}|\right)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U)$$
$$+\cdots \quad (28)$$

Finally, we can rewrite Eq. (1) according to the following substitutions of the equations provided above into Eq. (1): Eq. (15) is substituted into Eq. (1); Eq. (24) is substituted into $\overline{F}^{ex}[n]$ in Eq. (15) and Eq. (21) is substituted into $F_{HS}[n]$ in Eq. (15) ; Eq. (7) is substituted into $F_{IG}[n]$ in Eq. (21); Eq. (20) is substituted into $U_{IG}(\mathbf{r})$ in Eq. (7) and $n_{IG}\left(\mathbf{r}|U_{IG}\right)$ is replaced by $n(\mathbf{r}|U)$ because of Eqs. (16) and (23); Eq. (28) is substituted into $U_{HS}(\mathbf{r})$ in Eq. (20), resulting in

$$\Omega_{RMDFT}[U]=\Omega[0]-\frac{1}{\beta}\int d\mathbf{r}_1\Delta n(\mathbf{r}_1|U)+\Delta F_{HS}^{ex}[n]$$
$$-\int d\mathbf{r}_1\left[\frac{\delta F_{HS}^{ex}[n]}{\delta n(\mathbf{r}_1|U)}n(\mathbf{r}_1|U)-\mu_{HS}^{ex}n_0\right]$$
$$+\frac{n_0}{\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(2)}\left(|\mathbf{r}_1-\mathbf{r}_2|\right)\Delta n(\mathbf{r}_2|U)$$
$$+\frac{1}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(2)}\left(|\mathbf{r}_1-\mathbf{r}_2|\right)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U)$$
$$+\frac{n_0}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 C_{ex}^{(3)}\left(|\mathbf{r}_1-\mathbf{r}_2|,|\mathbf{r}_2-\mathbf{r}_3|,|\mathbf{r}_3-\mathbf{r}_1|\right)\Delta n(\mathbf{r}_2|U)\Delta n(\mathbf{r}_3|U)$$
$$+\frac{1}{3\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 C_{ex}^{(3)}\left(|\mathbf{r}_1-\mathbf{r}_2|,|\mathbf{r}_2-\mathbf{r}_3|,|\mathbf{r}_3-\mathbf{r}_1|\right)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U)\Delta n(\mathbf{r}_3|U)$$
$$+\cdots \quad , \quad (29)$$

where

$$\Omega[0]=F[n_0]-\mu N \quad , \quad (30)$$

and

$$\Delta F_{HS}^{ex}[n] = F_{HS}^{ex}[n] - F_{HS}^{ex}[n_0] \,. \tag{31}$$

In Eq. (32), we make a comparison of $\Omega[U]$s obtained using the RMDFT-type and HNC-type Taylor series expansions that are given by Eqs. (24) and (10), respectively:

$$\begin{aligned}
\Omega_{RMDFT}[U] - \Omega_{HNC}[U] = \Delta F_{HS}^{ex}[n] - \int d\mathbf{r}_1 \left[\frac{\delta F_{HS}^{ex}[n]}{\delta n(\mathbf{r}_1|U)} n(\mathbf{r}_1|U) - \mu_{HS}^{ex} n_0 \right] \\
- \frac{n_0}{\beta} \int d\mathbf{r}_1 d\mathbf{r}_2 C_{HS}^{(2)}(|\mathbf{r}_1 - \mathbf{r}_2|) \Delta n(\mathbf{r}_2|U) \\
- \frac{1}{2\beta} \int d\mathbf{r}_1 d\mathbf{r}_2 C_{HS}^{(2)}(|\mathbf{r}_1 - \mathbf{r}_2|) \Delta n(\mathbf{r}_1|U) \Delta n(\mathbf{r}_2|U) \\
- \frac{n_0}{2\beta} \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 C_{HS}^{(3)}(|\mathbf{r}_1 - \mathbf{r}_2|, |\mathbf{r}_2 - \mathbf{r}_3|, |\mathbf{r}_3 - \mathbf{r}_1|) \Delta n(\mathbf{r}_2|U) \Delta n(\mathbf{r}_3|U) \\
- \frac{1}{3\beta} \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 C_{HS}^{(3)}(|\mathbf{r}_1 - \mathbf{r}_2|, |\mathbf{r}_2 - \mathbf{r}_3|, |\mathbf{r}_3 - \mathbf{r}_1|) \Delta n(\mathbf{r}_1|U) \Delta n(\mathbf{r}_2|U) \Delta n(\mathbf{r}_3|U) \\
+ \cdots
\end{aligned}$$

.. (32)

The difference given by Eq. (32) suggests that the origin of the slow convergence in the HNC-type Taylor series expansion for dense liquids is attributed to the same origin that has been pointed out for dense HS fluids.[34,35] On the other hand, the difference vanishes at the low-density limit so that the HNC approximation gives the exact solution.

B. Integral equations for distribution functions

We can obtain an integral equation for the density distribution function under an arbitrary external field according to the following substitutions of the equations provided above: Eq. (23) is substituted into $n(\mathbf{r}|U) = n_{HS}(\mathbf{r}|U_{HS})$ in Eq. (16) and then Eq. (23) is substituted into $n_{HS}(\mathbf{r}|U_{HS})$ in that equation; Eq. (20) is substituted into $U_{IG}(\mathbf{r})$ in Eq. (5); Eq. (28) is substituted into $U_{HS}(\mathbf{r})$ in Eq. (20), resulting in

$$n(\mathbf{r}|U) = n_0 \exp[-\beta U_{IG}(\mathbf{r})] \,, \tag{33}$$

$$
\begin{aligned}
U_{IG}(\mathbf{r}) = U(\mathbf{r}) &- \frac{1}{\beta}\int d\mathbf{r}_1 C^{(2)}\left(|\mathbf{r}-\mathbf{r}_1|\right)\Delta n\left(\mathbf{r}_1|U\right) \\
&+ \frac{\delta F_{HS}^{ex}[n]}{\delta n\left(\mathbf{r}|U\right)} - \mu_{HS}^{ex} + \frac{1}{\beta}\int d\mathbf{r}_1 C_{HS}^{(2)}\left(|\mathbf{r}-\mathbf{r}_1|\right)\Delta n\left(\mathbf{r}_1|U\right) \\
&- \frac{1}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(3)}\left(|\mathbf{r}-\mathbf{r}_1|,|\mathbf{r}_1-\mathbf{r}_2|,|\mathbf{r}_2-\mathbf{r}|\right)\Delta n\left(\mathbf{r}_1|U\right)\Delta n\left(\mathbf{r}_2|U\right) \\
&+ \cdots
\end{aligned}
\tag{34}
$$

These equations represent some of the main results in this paper. Even if we neglect the $C_{ex}^{(3)}$ term and the higher-order terms in the Taylor series expansion, we can obtain the correction terms to the HNC approximation (the second line on the right-hand side of Eq. (34)) compared with Eq. (14). Not only the $C_{ex}^{(3)}$ term and the higher-order terms, but also that correction terms generated via the RMDFT-type Taylor series expansion, are regarded as the bridge function. Comparing Eq. (34) with Eq. (14) intuitively suggests that, for dense liquids, the convergence of the density functional Taylor series expansion introduced by the RMDFT using the HS reference system is faster than using the HNC as given by Eq. (14). In this work, we refer to the second- and third-order approximations in the RMDFT-type Taylor series expansion as RMDFT(D) and RMDFT(T), respectively.

First, we consider the second-order approximation in Eq. (34). Here, we define the second-order contribution to the bridge function:

$$
B^{(2)}(\mathbf{r}) = -\beta\left(\frac{\delta F_{HS}^{ex}[n]}{\delta n\left(\mathbf{r}|U\right)} - \mu_{HS}^{ex}\right) - \int d\mathbf{r}_1 C_{HS}^{(2)}\left(|\mathbf{r}-\mathbf{r}_1|\right)\Delta n\left(\mathbf{r}_1|U\right) .
\tag{35}
$$

If we combine Eqs. (33) and (34) and use Percus' relation (PR) [23,39], which is provided by Eq. (36), we can obtain a closed integral equation of $h^{(2)}(r)$ for a pure liquid.

$$
h^{(2)}(r) = n\left(r|U_{PR} = v_{vv}\right)/n_0 - 1 .
\tag{36}
$$

In Eq. (36), the pair correlation function for a pure liquid, $h^{(2)}(r)$, is exactly related to the density distribution function when the external field is chosen as the pair potential for the pure liquid, i.e., $U_{PR}(r) = v_{vv}(r)$. In this integral equation, $h^{(2)}(r)$, as well as the second-order direct correlation function, $C^{(2)}(r)$, can be determined in a self-consistent manner using the Ornstein-Zernike (OZ) equation:

$$
h^{(2)}\left(|\mathbf{r}-\mathbf{r}'|\right) = C^{(2)}\left(|\mathbf{r}-\mathbf{r}'|\right) + n_0\int d\mathbf{r}_1 C^{(2)}\left(|\mathbf{r}-\mathbf{r}_1|\right)h^{(2)}\left(|\mathbf{r}_1-\mathbf{r}'|\right) .
\tag{37}
$$

The OZ equation is derived from the following chain rule:

$$\int d\mathbf{r}_1 \left.\frac{\delta n(\mathbf{r}|U)}{\delta\left[-\beta U(\mathbf{r}_1)\right]}\right|_{U=0} \left.\frac{\delta\left[-\beta U(\mathbf{r}_1)\right]}{\delta n(\mathbf{r}'|U)}\right|_{n=n_0} = \delta\left(|\mathbf{r}-\mathbf{r}'|\right) \quad (38)$$

RMDFT(D) yields the integral equation for $h^{(2)}(r)$ that has been proposed by Rosenfeld as the reference HNC integral equation based on the universality of the ansatz for the bridge function,[30,40] if the Percus-Yevick solution for $C^{(2)}(r)$ and the excess free-energy functional model provided by the fundamental measure theory[35] are employed for the reference HS system. Hence, the RHNC equation can be systematically derived if an HS fluid is introduced as the reference system instead of an ideal gas and then the second-order truncation of the density-functional Taylor series expansion is applied to the excess part of the intrinsic free energy functional that is redefined by the RMDFT. In this RMDFT scheme, note that we assumed that the density distribution function under the arbitrary external field can be reproduced via that for the reference HS system by introducing an effective external field on the reference HS system [as illustrated in Fig. 1]. The derivation based on the RMDFT using the reference HS system implies that the RMDFT(D) integral equation would give a better description for $h^{(2)}(r)$ than the HNC approximation, especially for dense liquids.

In the case of an infinitely dilute solution, the pair correlation function between solute and solvent, $h_{uv}^{(2)}(r)$, is determined using Eqs. (33) and (34), and the following Percus' relation:

$$h_{uv}^{(2)}(r) = n\left(r|U_{PR}=v_{uv}\right)/n_0 - 1, \quad (39)$$

where $v_{uv}(r)$ is the solute-solvent interaction potential. In contrast to the integral equation for the pure liquid, $C^{(2)}(r)$ is used as the input in Eqs. (33) and (34).

Second, we consider the third-order approximation in Eq. (34), and define the third-order contribution to the bridge function:

$$B^{(3)}(\mathbf{r}) = \frac{1}{2}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(3)}\left(|\mathbf{r}-\mathbf{r}_1|,|\mathbf{r}_1-\mathbf{r}_2|,|\mathbf{r}_2-\mathbf{r}|\right)\Delta n(\mathbf{r}_1|U)\Delta n(\mathbf{r}_2|U). \quad (40)$$

A number of studies have been conducted to express the triplet direct correlation function based on factorization approximation.[41-45] In the same manner as the OZ equation of Eq. (37), $C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ is related to the triplet correlation

function, $h^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$, via the OZ3 equations that are derived from the functional derivatives of Eq. (38) [see APPENDIX A].[33] If we apply the following convolution approximation (CA) to $h^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$,[46,47]

$$\begin{aligned} h_{CA}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = & h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|)h^{(2)}(|\mathbf{r}_1-\mathbf{r}_3|) \\ & + h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|)h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|) \\ & + h^{(2)}(|\mathbf{r}_1-\mathbf{r}_3|)h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|) \\ & + n_0\int d\mathbf{r}_4 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_4|)h^{(2)}(|\mathbf{r}_2-\mathbf{r}_4|)h^{(2)}(|\mathbf{r}_3-\mathbf{r}_4|), \end{aligned} \tag{41}$$

the OZ3 equation yields $C_{CA}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)=0$ [see APPENDIX A]. This observation indicates that the convolution approximation leads to the HNC approximation given by Eq. (14). Iyetomi and Ichimaru proposed the following factorization approximation for $C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ [41,42,47]:

$$C_{CA}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|)h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|)h^{(2)}(|\mathbf{r}_3-\mathbf{r}_1|). \tag{42}$$

This factorization approximation yields the first correction to the convolution approximation for $h^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ and thus results in the following equation:

$$h_{CK}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = h_{CA}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) + \Delta h_{CK}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3), \tag{43}$$

where $\Delta h_{CK}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ is given by the OZ3 equation with Eq. (43) as a sum of $C_{CK}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ and the vertex correction terms of $C_{CK}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ [see APPENDIX B].

In the present study, we employ the factorization approximation shown in Eq. (42) to determine the third-order bridge function represented by Eq. (40).

III. CALCULATION DETAILS

A. Reference HS fluid

In the RMDFT calculation for dense simple liquids, an excess intrinsic free-energy functional, $F_{HS}^{ex}[n]$, and the second-order direct correlation function, $C_{HS}^{(2)}(r)$, for the reference HS fluid are needed. One can employ an arbitrary model for $F_{HS}^{ex}[n]$ and $C_{HS}^{(2)}(r)$ in the RMDFT. In this study, we employ the effective-density approximation (EDA)[38] for $F_{HS}^{ex}[n]$ and

$C_{HS}^{(2)}(r)$. If a multicomponent HS fluid is needed for the reference system in the RMDFT, one can use the free-energy functional model provided by the fundamental measure theory (FMT)[35,48] or its modified versions.[36,37] $F_{HS}^{ex}[n]$ is given by EDA as follows:

$$F_{HS}^{ex}[n] = \int d\mathbf{r}_1 n(\mathbf{r}_1|U) f_{HS}\left(n_{eff}(\mathbf{r}_1|U)\right), \tag{44}$$

where $n_{eff}(\mathbf{r}|U)$ is the effective density, which is assumed to be a functional of $n(\mathbf{r}|U)$.

A highly accurate $f_{HS}(n)$ is obtained from the Carnahan-Starling equation of state (Eqs. (45) and (46)).[23,49]

$$f_{HS}(n) = \frac{A_{HS}^{ex}}{N} = \frac{1}{\beta}\frac{\eta(4-3\eta)}{(1-\eta)^2}, \tag{45}$$

and

$$\eta = \pi n d_{HS}^3/6, \tag{46}$$

where d_{HS} is the diameter of the reference HS fluid. $n_{eff}(\mathbf{r}|U)$ is approximated by the first-order density-functional Taylor series expansion:

$$n_{eff}(\mathbf{r}|U) = n_0 + \int d\mathbf{r}_1 W(|\mathbf{r}-\mathbf{r}_1|)\left[n(\mathbf{r}_1|U) - n_0\right], \tag{45}$$

where the expansion coefficient, $W(r)$, is related to $C_{HS}^{(2)}(r)$ via

$$\hat{W}(k) = \left\{-2\beta f'_{HS}(n_0) + \sqrt{\left(2\beta f'_{HS}(n_0)\right)^2 - 4n_0 f''_{HS}(n_0)\hat{C}_{HS}(k)}\right\} \Big/ 2n_0 f''_{HS}(n_0). \tag{46}$$

In Eq. (46), $\hat{W}(k)$ and $\hat{C}_{HS}(k)$ are the Fourier transforms of $W(r)$ and $C_{HS}^{(2)}(r)$, respectively, and $f'_{HS}(n)$ and $f''_{HS}(n)$ are the first and second derivatives of $f_{HS}(n)$ with respect to n, respectively. $W(r)$ and $C_{HS}^{(2)}(r)$ are determined in a self-consistent manner by solving the OZ integral equation (Eq. (37)) using the following EDA closure:

$$h_{HS}(r) = n_{HS}\left(r\middle|U_{HS}^{PR} = v_{HS}\right)/n_0 - 1, \tag{47}$$

$$n_{HS}\left(r\middle|U_{HS}^{PR} = v_{HS}\right) = n_0 \exp\left[-\beta U_{IG}^{HS}(r)\right], \tag{48}$$

$$U_{IG}^{HS}(r) = v_{HS}(r) + \frac{\delta F_{HS}^{ex}[n_{HS}]}{\delta n_{HS}\left(r\middle|U_{HS}^{PR}\right)} - \mu_{HS}^{ex}\,, \tag{49}$$

$$\frac{\delta F_{HS}^{ex}[n_{HS}]}{\delta n_{HS}\left(r\middle|U_{HS}\right)} = f_{HS}\left(n_{HS}\left(r\middle|U_{HS}\right)\right) + \int d\mathbf{r}_1 W\left(\left|\mathbf{r}-\mathbf{r}_1\right|\right) n_{HS}\left(\mathbf{r}_1\middle|U_{HS}\right) f_{HS}\left(n_{HS}\left(\mathbf{r}_1\middle|U_{HS}\right)\right), \tag{50a}$$

$$\mu_{HS}^{ex} = f_{HS}(n_0) + n_0 \int d\mathbf{r}_1 W\left(\left|\mathbf{r}-\mathbf{r}_1\right|\right) f_{HS}(n_0). \tag{50b}$$

Equation (47) is Percus' relation[23,39] where $U_{HS}^{PR}(r)$ is equal to the interaction potential between the reference HS particles, $v_{HS}(r)$.

B. Pair correlation functions for pure solvent and solute-solvent systems

From Eqs. (33)-(36), (40), and (42), the closure relations for pure solvent and for solute-solvent system are summarized as follows:

$$h^{(2)}(r) = n\left(r\middle|U_{PR} = v_{vv}\right)/n_0 - 1 \qquad \text{(pure solvent system)}, \tag{51a}$$

$$h_{uv}^{(2)}(r) = n\left(r\middle|U_{PR} = v_{uv}\right)/n_0 - 1 \qquad \text{(solute-solvent system)}, \tag{51b}$$

$$n\left(r\middle|U_{PR}\right) = n_0 \exp\left[-\beta U_{IG}(r)\right], \tag{52}$$

$$U_{IG}(r) = v_{vv}(r) - \frac{1}{\beta}\left\{h^{(2)}(r) - C^{(2)}(r)\right\} - \frac{1}{\beta}\left\{B^{(2)}(r) - B_{CK}^{(3)}(r)\right\} \qquad \text{(pure solvent system)}, \tag{53a}$$

$$U_{IG}(r) = v_{uv}(r) - \frac{1}{\beta}\int d\mathbf{r}_1 C^{(2)}\left(\left|\mathbf{r}-\mathbf{r}_1\right|\right)\Delta n\left(\mathbf{r}_1\middle|U_{PR}\right) - \frac{1}{\beta}\left\{B^{(2)}(r) + B_{CK}^{(2)}(r)\right\} \qquad \text{(solute-solvent system)}, \tag{53b}$$

$$B^{(2)}(\mathbf{r}) = -\beta\left(\frac{\delta F_{HS}^{ex}[n]}{\delta n\left(\mathbf{r}\middle|U_{PR}\right)} - \mu_{HS}^{ex}\right) - \int d\mathbf{r}_1 C_{HS}^{(2)}\left(\left|\mathbf{r}-\mathbf{r}_1\right|\right)\Delta n\left(\mathbf{r}_1\middle|U_{PR}\right)\,, \tag{54}$$

$$B_{CK}^{(3)}(\mathbf{r}) = \frac{1}{2}\int d\mathbf{r}_1\, d\mathbf{r}_2 h^{(2)}\left(\left|\mathbf{r}-\mathbf{r}_1\right|\right) h^{(2)}\left(\left|\mathbf{r}_1-\mathbf{r}_2\right|\right) h^{(2)}\left(\left|\mathbf{r}_2-\mathbf{r}\right|\right)\Delta n\left(\mathbf{r}_1\middle|U_{PR}\right)\Delta n\left(\mathbf{r}_2\middle|U_{PR}\right). \tag{55}$$

Equation (51a) (or (51b)) is Percus' relation [23,39] where $U_{PR}(r)$ is equal to the interaction potential between solvent particles, $v_{vv}(r)$ (or between solute and solvent particles $v_{uv}(r)$). In Eq. (54), the first term on the right-hand side is given by Eqs. (50a) and (50b).

Here we explain the detailed derivation of Eq. (55) from Eq. (40). In Eq. (40), for $B^{(3)}(r)$, the triplet direct correlation function is given by $C_{ex}^{(3)}$, i.e., $C^{(2)}-C_{HS}^{(2)}$, as defined by Eq. (27). In our preliminary calculation, we applied the factorization approximation of Eq. (42) to both $C^{(3)}$ and $C_{HS}^{(3)}$ in order to calculate $B^{(3)}(r)$ using Eq. (40). However, we obtained neither effective bridge function corrections nor a small contribution to the SFE, discussed in the next subsection. Therefore, we replaced $C_{ex}^{(3)}$ with $C^{(3)}$ in Eq. (40), and then applied the factorization approximation of Eq. (42) to $C^{(3)}$, so that $C_{ex}^{(3)} \approx C_{CK}^{(3)}$. As a result, we obtained Eq. (55). From a theoretical point of view, the assumption $C_{ex}^{(3)} \approx C^{(3)}$ is valid if the free-energy functional model, $F_{HS}^{ex}[n]$, for the reference HS system contains no effective contribution originating from the $C_{HS}^{(3)}$-term. We numerically assesse the validity of the assumption, $C_{ex}^{(3)} \approx C_{CK}^{(3)}$, by comparing results calculated using $B_{CK}^{(3)}(r)$ with those provided by MD simulation.

C. The SFE functional

The SFE functional in the infinite dilute limit of the solute is given by $\Delta G_{solv} = \Omega[U_{PR} = v_{uv}] - \Omega[0]$. Using Eqs. (29), (30), (50a), and (50b), we obtained the SFE functional as follows:

$$\begin{aligned}\Delta G_{solv} = &-\frac{1}{\beta}\int d\mathbf{r}_1 \Delta n(\mathbf{r}_1|U_{PR}) - \int d\mathbf{r}_1 \Big[n(\mathbf{r}_1|U_{PR}) \Big\{ \int d\mathbf{r}_2 W(|\mathbf{r}_1-\mathbf{r}_2|) \Big[n(\mathbf{r}_2|U_{PR}) f_{HS}'(n_{eff}(\mathbf{r}_2|n)) - n_0 f_{HS}'(n_0) \Big] \\ &+ \int d\mathbf{r}_1 W(|\mathbf{r}_1|) n_0 f_{HS}'(n_0) \Big\} - n_0 f_{HS}'(n_0) \int d\mathbf{r}_1 W(|\mathbf{r}_1|) n_0 \Big] + \frac{n_0}{\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|) \Delta n(\mathbf{r}_2|U_{PR}) \\ &+ \frac{1}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2 C_{ex}^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|) \Delta n(\mathbf{r}_1|U_{PR}) \Delta n(\mathbf{r}_2|U_{PR}) + \Delta G_{CK}^{(3)}\end{aligned} \quad ,(56)$$

$$\begin{aligned}\Delta G_{CK}^{(3)} = &\frac{n_0}{2\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|) h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|) h^{(2)}(|\mathbf{r}_3-\mathbf{r}_1|) \Delta n(\mathbf{r}_2|U_{PR}) \Delta n(\mathbf{r}_3|U_{PR}) \\ &+ \frac{1}{3\beta}\int d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|) h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|) h^{(2)}(|\mathbf{r}_3-\mathbf{r}_1|) \Delta n(\mathbf{r}_1|U_{PR}) \Delta n(\mathbf{r}_2|U_{PR}) \Delta n(\mathbf{r}_3|U_{PR})\end{aligned} \quad , \quad (57)$$

$$U_{PR}(r) = v_{uv}(r). \quad (58)$$

In Eq. (56), $\Delta G_{CK}^{(3)}$ is neglected in RMDFT(D), while in the case of RMDFT(T), $\Delta G_{CK}^{(3)}$, given by Eq. (57), is taken into account.

D. Models for pure liquid and solute-solvent systems

In this study, we apply the RMDFT integral equations to the Lennard-Jones (LJ) pure and infinitely-dilute LJ solution systems. The interaction potential between LJ particles is given by

$$v_{ij}(r) = 4\varepsilon_{ij}\left[\left(d_{ij}/r\right)^{12} - \left(d_{ij}/r\right)^{6}\right], \tag{59a}$$

$$d_{ij} = \left(d_i + d_j\right)/2, \tag{59b}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}, \tag{59c}$$

where ε_i and d_i are the parameters for the potential energy depth and diameter, respectively. The interaction potentials for solvent particles, and for solute-solvent particles are defined as $v_{vv}(r)$ and $v_{uv}(r)$, respectively. In order to determine the diameter for the reference HS fluid in the RMDFT, we applied the RMDFT integral equation to a LJ liquid near the triple point ($n_0 d_v^3 = 0.84$ and $T/\varepsilon_v = 0.75$), and then employed $1.013 d_v$ as the optimal diameter for the reference HS fluid based on comparison with results obtained using the RMDFT integral equation and those provided by a MD simulation.[50] Computational time needed to perform the RMDFT(D) or the RMDFT(T) for the LJ liquid near the triple point using the single core of 1.7 GHz Intel Core i7 processor is about 6 seconds or 1010 minutes. Using this HS diameter determined based on the bulk properties, we applied the RMDFT integral equation to the infinitely dilute LJ solutions at a different thermodynamic state ($n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$) where MD simulation results for the SFE and pair distribution functions have been reported.[27] Here, it is noted that the SFE calculated using the RMDFT is sensitive to the diameter of the reference HS system. In our previous study, we found that the values of SFE provided by the RMDFT monotonically decrease as the diameter of the reference HS system increases. [28,29] Thus, we determined the optimal HS diameter on the basis of experimental values of SFE for methane, propane, and isobutane. In general, the optimal HS diameter would depend on the thermodynamic condition. The systematic way for determination of the reference HS diameter is one of the important issues that are needed to apply the RMDFT to various systems.

IV. RESULTS AND DISCUSSION

A. Bulk Lennard-Jones liquid near the triple point

It is well known that the HNC approximation overestimates the height of the first peak in the pair distribution function, $g^{(2)}(r)$, underestimates the radial distance of the first peak position in $g^{(2)}(r)$, and overestimates the isothermal compressibility, χ_T, for LJ liquids.[30] These theoretical drawbacks in the HNC approximation can be observed especially at a

thermodynamic state near the triplet point. Figure 2 (a) shows $g^{(2)}(r)$ obtained using three approximations from the DFT (i.e., HNC, RMDFT(D), and RMDFT(T)) and a MD simulation [50] for an LJ liquid at a thermodynamic state near the triplet point of $n_0 d_v^3 = 0.84$ and $T/\varepsilon_v = 0.75$. Computational details for solving the integral equations are given in APPENDIX C and D. In Fig. 2 (a), we see a shift of the first peak position toward a shorter distance in $g^{(2)}(r)$ obtained using the HNC approximation rather than using the MD simulation. This drawback of the HNC approximation for $g^{(2)}(r)$ is improved by the second-order bridge function, $B^{(2)}(r)$, i.e., RMDFT(D) as well as the second-order plus third-order bridge functions, $B^{(2)}(r)+B_{CK}^{(3)}(r)$, i.e., RMDFT(T). $B^{(2)}(r)$ mainly affects the first peak of $g^{(2)}(r)$, whereas the result for RMDFT(D) almost overlaps with the result by HNC at distances of r/d_v>1.4. These observations are also shown in Fig. 2 (b) by the difference between $g^{(2)}(r)$s, $\Delta g^{(2)}(r) = g_{\text{theo}}^{(2)}(r) - g_{\text{sim}}^{(2)}(r)$, where the subscripts "theo" and "sim" refer to the theoretically calculated results from the present work and the available simulation results, [50] respectively. As seen in Figs 2 (a) and 2 (b), the third-order bridge function, $B_{CK}^{(3)}(r)$, can improve the shift of the first peak position in $g^{(2)}(r)$ and the underestimation of $g^{(2)}(r)$ at distances longer than the first maximum ($1.1 < r/d_v < 1.4$), and the overestimation of $g^{(2)}(r)$ at distances around the first minimum ($1.4 < r/d_v < 1.8$). Figure 2 (b) clearly shows the systematic improvements made by the second- and third-order bridge functions. The larger reduction in $\Delta g^{(2)}(r)$ is attained by taking into account $B_{CK}^{(3)}(r)$. Similar effects of $B_{CK}^{(3)}(r)$ on the second peak of $g^{(2)}(r)$, as seen in the result by RMDFT(T), i.e., the broadening of the second peak, has also been observed in the case of one component plasmas with the coupling constant, Γ, having values larger than 200.[41,47] Although the applicability of the bridge function based on Eqs. (40) and (42) to fluids with long-range Coulomb interactions has been demonstrated, [41,47] to the best of our knowledge, we show for the first time that the factorization approximation based on Eq. (42) is effective in constructing a bridge function for liquids with non-Coulomb interactions if the third-order bridge function, $B_{CK}^{(3)}(r)$, and the second-order bridge function, $B^{(2)}(r)$, are simultaneously applied. The detailed shape of these bridge functions, including $B_{CK}^{(3)}(r)$, will be discussed in Fig. 3.

In Fig. 2 (c), we see a large difference between the HNC and MD results with respect to the second-order direct correlation function, $C^{(2)}(r)$. The overestimation of $C^{(2)}(r)$ by HNC at distances of $r/d_v < 1$ is improved greatly by

RMDFT(D). In addition, as shown in Fig. 2 (d) by the difference between $C^{(2)}(r)$s, $\Delta C^{(2)}(r) = C^{(2)}_{\text{theo}}(r) - C^{(2)}_{\text{sim}}(r)$, we can see the further improvement of $C^{(2)}(r)$ by RMDFT(T). The accuracy of $C^{(2)}(r)$ for homogeneous liquid is important in RMDFT and HNC to reproduce the density distribution function for inhomogeneous fluids under an external field, since $C^{(2)}(r)$ is required as the second-order coefficient in the density-functional Taylor series expansions.

Figures 2 (e) and 2 (f) show the structure factors, $S(k)$, and their differences, $\Delta S(k) = S_{\text{theo}}(k) - S_{\text{sim}}(k)$, respectively. $S(k)$ is an important bulk property, because $S(k=0)$ is related to the isothermal compressibility via $\chi_T = S(k=0)/n_0 k_B T$. $S(k)$ also corresponds to the density response function and, thus, can be regarded as an indicator for whether the corresponding $C^{(2)}(r)$ term is valid as the expansion coefficient in the density-functional Taylor series expansion. As seen in Fig. 2 (e) and Fig. 2 (f), the drawback of the HNC approximation for $S(k)$ in the small *k*-region, i.e., the large overestimation of the $S(k \sim 0)$ value, is improved by $B^{(2)}(r)$. However, $B^{(3)}_{CK}(r)$ provides no obvious effect on $S(k)$ in the small *k*-region. On the other hand, we find that $B^{(2)}(r)$ does not significantly affect $S(k)$ for $r/d_v > 7$, and thus the result by RMDFT(D) almost overlaps the result by HNC at $r/d_v > 7$, while $B^{(3)}_{CK}(r)$ improves the underestimation of $S(k)$ at the first peak, the overestimation of $S(k)$ at the *k*-values around the first minimum ($8 < r/d_v < 10$), and the detailed shape of $S(k)$ around the second peak ($10 < r/d_v < 13$). Figure 2 (f) clearly shows the systematic improvements on $S(k)$ by the second- and third-order bridge functions.

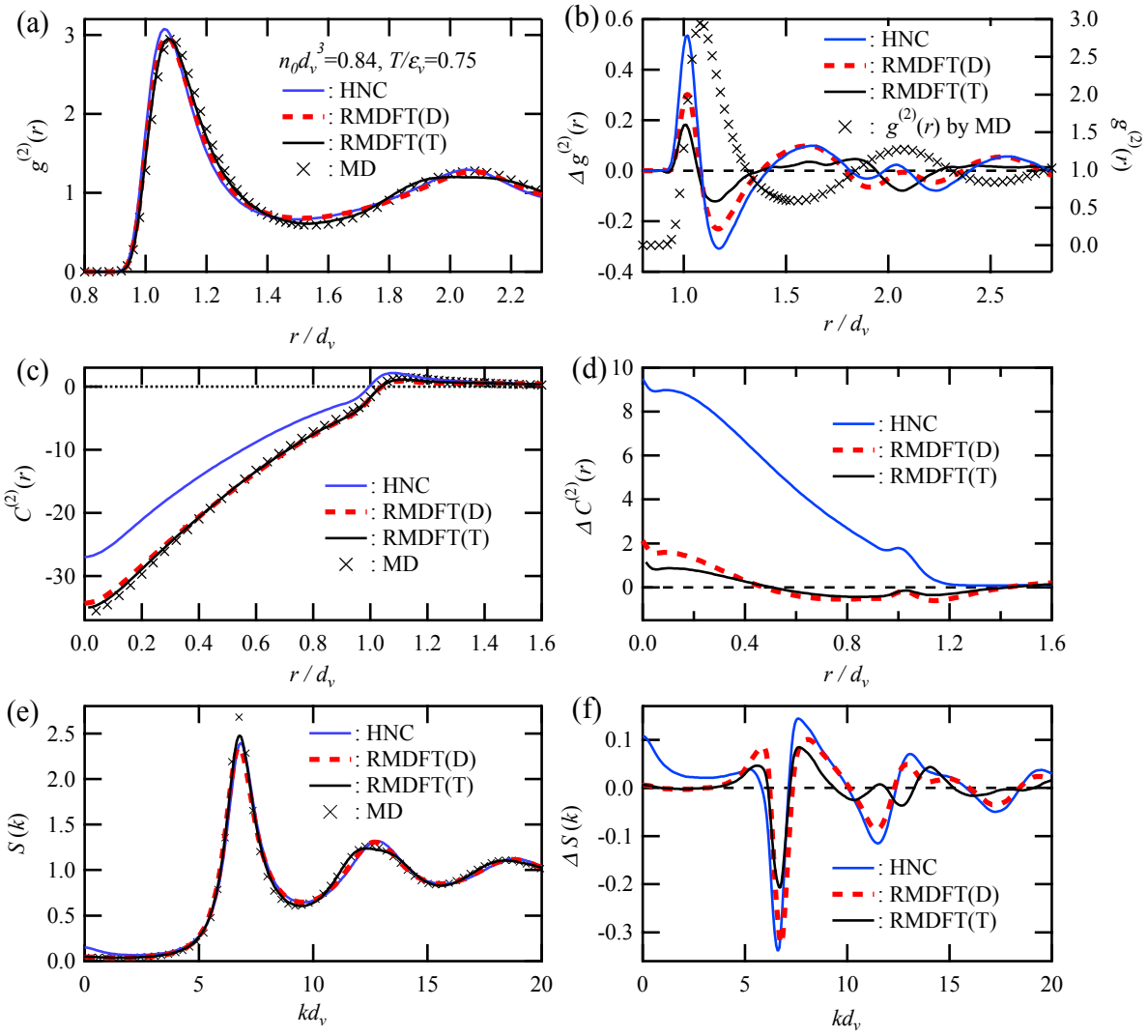


FIG. 2 Comparison of results obtained using three approximations from the DFT and a molecular dynamics simulation[50] for a Lennard-Jones liquid near the triple point ($n_0 d_v^3 = 0.84$ and $T/\varepsilon_v = 0.75$). (a) Pair distribution function, $g^{(2)}(r)$, (b) difference between the theoretical and simulation results of $g^{(2)}(r)$, $\Delta g^{(2)}(r) = g^{(2)}_{\text{theo}}(r) - g^{(2)}_{\text{sim}}(r)$, (c) second-order direct correlation function, $C^{(2)}(r)$, (d) difference between the theoretical and simulation results for $C^{(2)}(r)$, $\Delta C^{(2)}(r) = C^{(2)}_{\text{theo}}(r) - C^{(2)}_{\text{sim}}(r)$, (e) structure factor $S(k)$, and (f) difference between the theoretical and simulation results for $S(k)$, $\Delta S(k) = S_{\text{theo}}(k) - S_{\text{sim}}(k)$. RMDFT(T) and RMDFT(D) indicate the results obtained using Eqs. (53a) and (54), respectively, with and without the third-order bridge function given by Eq. (55), while HNC indicates results obtained using Eq. (53a) without these bridge functions given by Eqs. (54) and (55).

In Fig. 3 (a), the shapes of $B^{(2)}(r)$ and $B^{(3)}_{CK}(r)$ are shown for comparison with the bridge functions that have been discussed in the literatures.[42,44] We find that both $B^{(2)}(r)$ and $B^{(3)}_{CK}(r)$ have negative values at distances smaller than $r/d_v \sim 1$. The negative values of $B^{(2)}(r)$ around $r/d_v = 1$ correspond to effective repulsions between the particles, so that $B^{(2)}(r)$ suppresses the height of the first peak in $g^{(2)}(r)$, and pushes the position of its first peak toward longer distances. We also find that the contribution from the third-order term, $B^{(3)}_{CK}(r)$, to the total bridge function is smaller than

that from the second-order term, $B^{(2)}(r)$. In fact, the total bridge function does not seem to differ very much from $B^{(2)}(r)$ in the scale of Fig. 3(a). The relatively smaller contribution of $B_{CK}^{(3)}(r)$ is consistent with the general nature of the Taylor series in the case where it converges.

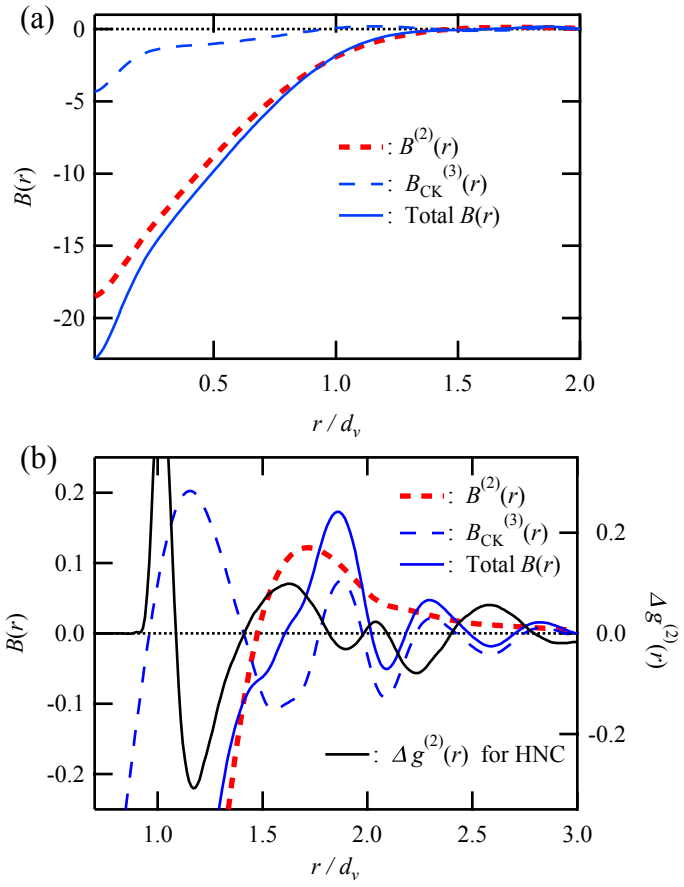


FIG. 3 (a) The second- and third-order bridge functions given, respectively, by Eqs. (54) and (55) and the total bridge function as the sum of them. (b) These bridge functions and $\Delta g^{(2)}(r)$ are shown on the left and right axis, respectively.

In Fig. 3 (b), the shape of the bridge functions at distances between the first and third peaks for $C^{(2)}(r)$ is shown together with $\Delta g^{(2)}(r)$ for the HNC approximation. We find that $B^{(2)}(r)$ represents short-range repulsion between LJ particles because $B^{(2)}(r)$ has large negative values at distances around the first peak of $g^{(2)}(r)$. On the other hand, $B_{CK}^{(3)}(r)$ provides a positive maximum in the vicinity of the first peak for $g^{(2)}(r)$, so that $B_{CK}^{(3)}(r)$ improves the underestimation of $g^{(2)}(r)$ by either HNC or RMDFT(D) at distances where $\Delta g^{(2)}(r)$ has negative values ($1.1 < r/d_v <$

1.4). In addition to the effects of $B_{CK}^{(3)}(r)$ on the shape of the first peak in $g^{(2)}(r)$, $B_{CK}^{(3)}(r)$ improves the overestimation of $g^{(2)}(r)$ around the first minimum ($1.4 < r/d_v < 1.8$) [see Fig. 2 (a)], and also results in broadening the second peak of $g^{(2)}(r)$, since $B_{CK}^{(3)}(r)$ has two small positive maximums at distances slightly shorter and longer than the second peak position of $g^{(2)}(r)$. It should be noted that $B_{CK}^{(3)}(r)$ and $\Delta g^{(2)}(r)$ have oscillation structures with opposite signs at distances longer than $r/d_v = 1.1$. Thus, $B_{CK}^{(3)}(r)$ can yield appropriate bridge corrections to the HNC approximation.

In a previous study, Iyetomi and Ichimaru introduced a renormalization function, $\phi(r)$, to construct a bridge function such as $\phi(r)B_{CK}^{(3)}(r)$, and then improved the HNC approximation for dense one-component plasmas.[42] In their scheme, $B_{CK}^{(3)}(r)$ is enlarged by $\phi(r)$, since $B_{CK}^{(3)}(r)$ cannot provide large enough corrections. As a result, the renormalized bridge function, $\phi(r)B_{CK}^{(3)}(r)$, provides appropriate improvements in dense one-component plasmas. For $g^{(2)}(r)$, the underestimation of the first peak, the overestimation of the first minimum, and the underestimation around the second peak provided by the HNC approximation are improved by $\phi(r)B_{CK}^{(3)}(r)$. However, since $B_{CK}^{(3)}(r)$ provides a positive maximum at distances around the first peak of $g^{(2)}(r)$, it is easy to expect that $B_{CK}^{(3)}(r)$ would not give appropriate bridge corrections for hard-sphere fluids and LJ liquids. The observation would be related to the behavior of Eq. (42) in the short-range limit, i.e., $C_{CK}^{(3)} \sim -1$, due to $h^{(2)}(r \to 0) = -1$. [43,44] In contrast to the HNC-type Taylor series expansion given by Eq. (10) or (14), by introducing an appropriate HS fluid as the reference system instead of an ideal gas, the RMDFT-type Taylor series expansion yields not only $B_{CK}^{(3)}(r)$ but also $B^{(2)}(r)$, where the latter mainly works at the distances smaller than $r/d_v = 1.4$ as the most essential repulsive bridge correction [see Fig. 2 (b)], while the former assisted by $B^{(2)}(r)$ can provide the further corrections not only for the small distances but also for the distances larger than $r/d_v = 1.4$ [also see Fig. 2 (b)].

B. Infinitely dilute Lennard-Jones solutions

In section IVA, we investigated the validity of the RMDFT integral equations for the LJ liquid, and demonstrated an improvement in the pair and direct correlation functions and the structure factor. These correlation functions for inhomogeneous liquids are required as inputs to solve the RMDFT integral equations for inhomogeneous liquids, such as

infinitely dilute solutions. In this subsection, we apply the RMDFT integral equations to infinitely dilute LJ solutions at a thermodynamic state of $n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$ where MD simulation results have been reported[27] in order to investigate the validity of the RMDFT for the solute-solvent systems. The LJ potential between the solute and solvent particles is given by Eqs. (59a-c), in which the diameter of the solute, d_u, is chosen for the three values, d_v, $2d_v$, and $3d_v$, while the energy parameter for the solute ε_u is always chosen to be ε_v.

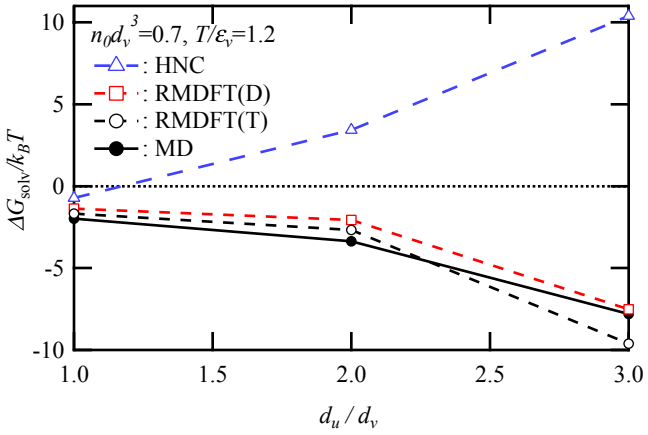


FIG. 4 Comparison of SFE values for a LJ solute in a LJ solvent ($n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$) obtained using three approximations from the DFT and MD simulations.[27] Results for RMDFT(T) and RMDFT(D) are obtained using Eq. (56) with and without the third-order bridge function given by Eq. (57), respectively. The statistical errors in MD simulations are less than the size of the corresponding symbol.

Figure 4 shows the solute size dependence of SFE for a LJ solute in an LJ solvent. The SFE value when $d_u = d_v$ corresponds to the excess chemical potential for the LJ pure liquid. In this case, the HNC approximation provides a SFE value relatively close to the result obtained by the MD simulation. [27] At the thermodynamic state examined in this study, the SFE value calculated by the MD simulation monotonically decreases as the size of the solute increases.[27] It is noted, however, that the solute size dependence of SFE depends not only on the thermodynamic state but also on the solute-solvent energy parameter, ε_{uv}.[51] In contrast to the result obtained by the MD simulation, the SFE value determined by HNC increases as the size of the solute increases. As a result, HNC produces a large value of SFE for large solutes. On the other hand, RMDFT(D) and RMDFT(T) greatly improve the overestimation of SFE, thus yielding good agreement with the result obtained by the MD simulation, although RMDFT(T) slightly underestimates the value of SFE in the case where $d_u = 3d_v$.

Before we discuss the validity of $B^{(2)}(r)$ and $B_{CK}^{(3)}(r)$ in the solute-solvent pair distribution functions, we show how $B^{(2)}(r)$ and $B_{CK}^{(3)}(r)$ improve $g^{(2)}(r)$ for the pure LJ solvent at the thermodynamic state of $n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$. Figure 5(a) shows the comparison of $g^{(2)}(r)$s obtained using three approximations from the DFT and the MD simulation for the pure LJ solvent. The shift of the first peak for $g^{(2)}(r)$ toward a shorter distance obtained by HNC is improved by RMDFT(D) and RMDFT(T). However, RMDFT(D) and HNC underestimate $g^{(2)}(r)$ at distances longer than the first maxima ($1.2 < r/d_v < 1.5$). This underestimation is improved by RMDFT(T), because $B_{CK}^{(3)}(r)$ has a positive maximum at distances around the first peak, as shown in Fig. 5 (b). The range of the left axis in Fig. 5 (b) is chosen to be smaller than that in Fig. 3 (b) in order to show the detailed shape of the bridge functions. Thus, the effect of $B_{CK}^{(3)}(r)$ on $g^{(2)}(r)$ at this thermodynamic state is smaller than that near the triple point. Based on all the pure solvent results provided by RMDFT and the MD simulation, we conclude as follows. RMDFT(D) with the second-order bridge function, $B^{(2)}(r)$, works well for bulk properties such as $g^{(2)}(r)$ and excess chemical potential. Furthermore, the third-order bridge function, $B_{CK}^{(3)}(r)$, based on the factorization approximation provided by Eq. (42), yields further corrections for the LJ liquids if not only $B_{CK}^{(3)}(r)$ but also $B^{(2)}(r)$ are taken into account, although using only $B_{CK}^{(3)}(r)$ causes deterioration of the accuracy for $g^{(2)}(r)$ in the LJ liquids.

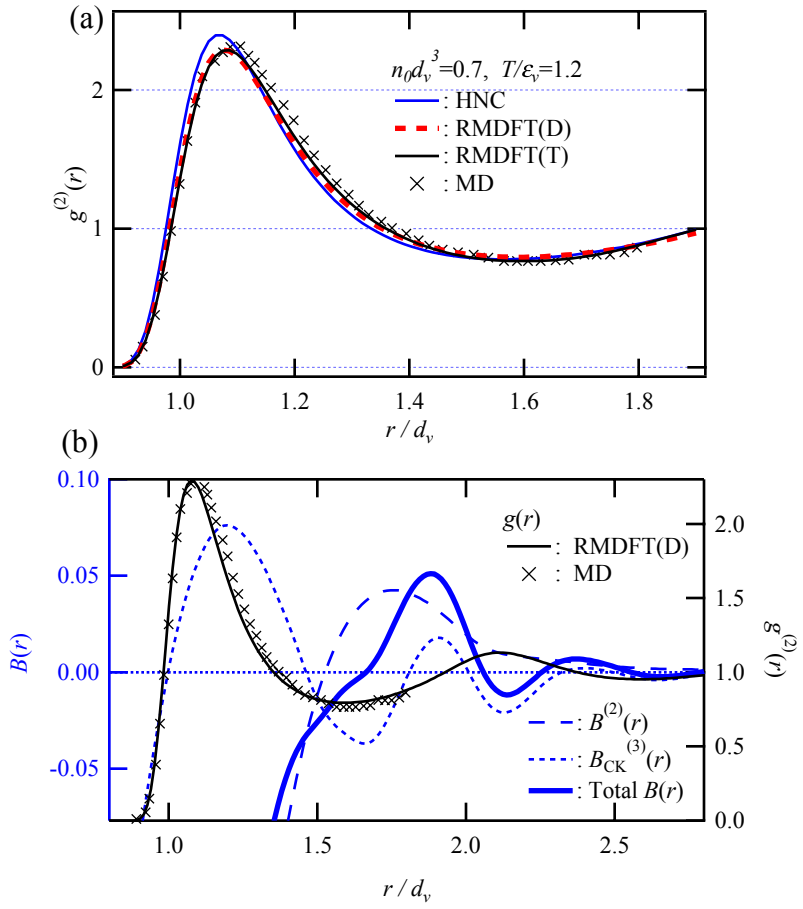


FIG. 5 (a) Comparison of the pair distribution functions for the LJ solvent ($n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$) obtained using three approximations from the DFT and a MD simulation.[27] (b) The bridge functions and pair distribution functions for the LJ solvent are indicated on the left and right axis, respectively.

Figure 6 (a) shows $g^{(2)}(r)$s obtained using three approximation from the DFT and the MD simulation for the solute-solvent system with $d_u = 3d_v$ where the HNC approximation significantly overestimates the SFE value [see Fig. 4]. In Fig. 6 (b), the bridge functions are shown on the left axis, and $g^{(2)}(r)$ is shown on the right axis for comparison with the detailed shape of the bridge functions. The first peak determined by the HNC approximation is remarkably shifted toward shorter distances. The well-known drawback of HNC on $g^{(2)}(r)$ is significantly increased by an increase in the size of the solute. As pointed out by Miyata and Thapa, a large overestimation of the SFE for large solute is yielded by the thermodynamic integral based on $g^{(2)}(r)$ calculated using the HNC closure because of the large shift in the first peak of $g^{(2)}(r)$.[27] RMDFT(D) and RMDFT(T) can remedy the shift of the first peak in $g^{(2)}(r)$, but slightly underestimate the height of the

first peak. The effect of $B_{CK}^{(3)}(r)$ on $g^{(2)}(r)$ is smaller than that of $B^{(2)}(r)$, so that the total bridge function, $B^{(2)}(r)+B_{CK}^{(3)}(r)$, almost overlaps $B^{(2)}(r)$. However, $B_{CK}^{(3)}(r)$ slightly reduces the height of the first peak of $g^{(2)}(r)$, since $B_{CK}^{(3)}(r)$ has a small negative minimum at a distance near the first peak of $g^{(2)}(r)$. In Fig. 6 (b), $B_{CK}^{(3)}(r)$ for the pure solvent, which has already been shown in Fig. 5(b), is displayed in order to compare $B_{CK}^{(3)}(r)$ for the solute-solvent system with $d_u = 3d_v$. The oscillation behavior of $B_{CK}^{(3)}(r)$ for the solute-solvent system is similar to that for the pure solvent system. The slight underestimation of the SFE value for the solute with $d_u = 3d_v$ might be related to the underestimation of the height of the first peak in the solute-solvent $g^{(2)}(r)$. As discussed in the APPENDIX D, the third-order bridge function $B_{CK}^{(3)}(r)$ with the factorization approximation for $C^{(3)}$ gives the simplest elementary bridge diagram with a $h^{(2)}$-bond. Therefore, in the theoretical point of view, $B_{CK}^{(3)}(r)$ seems to be valid for, at least, one-component homogeneous liquids. As for the numerical results for the bulk LJ liquids, the excess chemical potential [shown by Fig. 4] as well as $g^{(2)}(r)$ [shown by Fig. 5] is obviously improved by the third-order bridge function. These observations suggest that $B_{CK}^{(3)}(r)$ based on the factorization approximation provides an appropriate bridge correction for bulk LJ liquids, while it would not work as the effective bridge function for inhomogeneous systems such as the solute-solvent systems in the case where the size of the solute is larger than or different from the size of the solvent. We suppose that the use of the factorization approximation for $C^{(3)}$ in the third-order bridge function would be effective for only one-component homogeneous liquids.

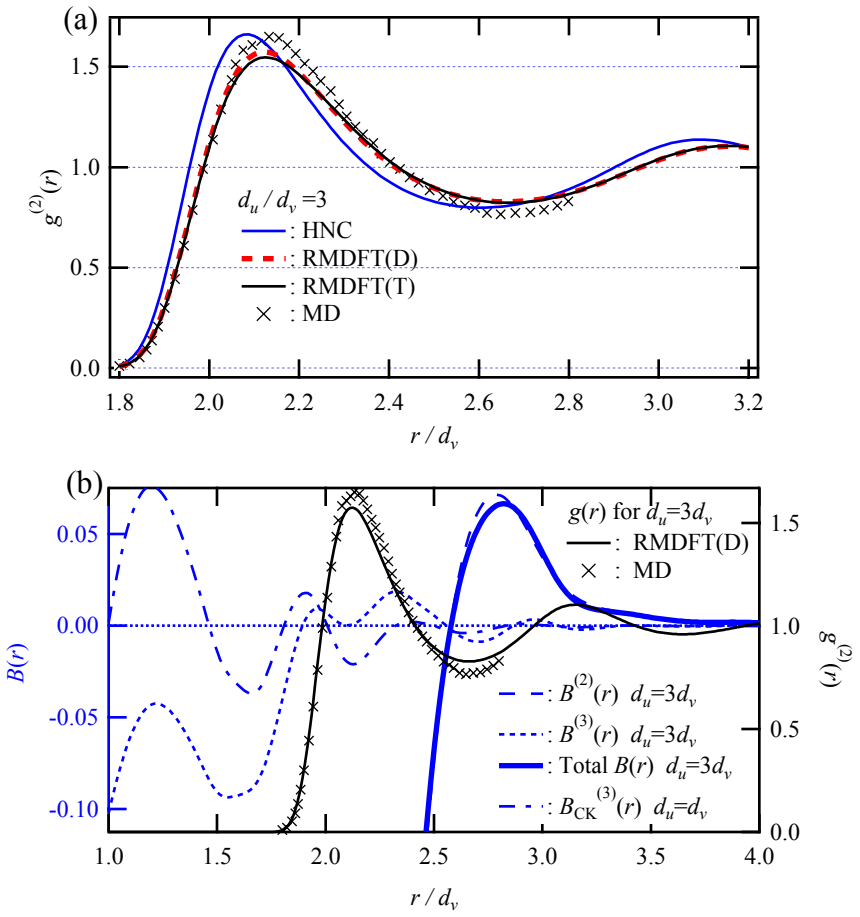


FIG. 6 (a) Comparison of the pair distribution functions between solute and solvent ($n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$) obtained using three approximations from the DFT and a MD simulation[27] in the case where the size of the solute is equal to $d_u/d_v = 3$. (b) The bridge and pair distribution functions for the same solute-solvent system are indicated by the left and right axis, respectively. A comparison with the third-order bridge function, $B_{CK}^{(3)}(r)$, for the solute-solvent system, and $B_{CK}^{(3)}(r)$ for the solvent system, i.e., $d_u = d_v$, are also shown in Fig. 6 (b).

V. CONCLUSION

In the present study, we propose a reference-modified density functional theory (RMDFT) as a systematic approach to improve both the SFE functional and the pair distribution functions for dense liquids. In order to assess the validity of the RMDFT approach, we applied RMDFT to a Lennard-Jones (LJ) liquid and infinitely dilute LJ solutions. In this approach, a hard-sphere fluid is introduced as the reference system in the DFT instead of an ideal gas that is normally used as the standard reference system. The contribution to the free energy caused by an excess interaction potential, which is defined as the difference of the interaction potentials between the system of interest and the reference HS system, is taken into account via the density functional Taylor series expansion of the excess intrinsic free-energy functional, which is redefined by

RMDFT as the difference of the intrinsic free energies between the system of interest and the reference HS system. The RMDFT-type density-functional Taylor series expansion theoretically yields more rapid convergence for dense liquids than the conventional HNC-type Taylor series expansion where an ideal gas is employed as the reference system. The free-energy functional obtained using RMDFT suggests that the problem of the slow convergence of the HNC-type series expansion for a dense liquid is essentially the same problem as the slow convergence of the HNC-type series expansion for an HS fluid, which has been pointed out previously.[34,35] The convergence of the HNC-type series expansion for a dense liquid would, therefore, be as slow as that of a dense HS fluid.

In the present study, we examined the third-order Taylor series expansion, as well as the second-order expansion, to calculate the pair distribution functions and excess chemical potential, or the solvation free energy (SFE). A factorization approximation for a triplet direct correlation function [41,42,47] was applied to evaluate the third-order term. For LJ liquids, the second-order approximation, namely RMDFT(D), improved the pair distribution function, second-order direct correlation function, structure factor, and excess chemical potential. In addition, further improvements of these bulk properties were attained by the third-order approximation, namely, RMDFT(T). To the best of our knowledge, this is the first time that a third-order bridge function based on the factorization approximation, represented by Eq. (42), is shown to yield appropriate bridge corrections for a non-Coulomb system, such as LJ liquid.

For infinitely dilute LJ solutions at the thermodynamic state of $n_0 d_v^3 = 0.7$ and $T/\varepsilon_v = 1.2$, MD simulations give a decrease in SFE as the size of the solute increases, while the HNC approximation predicts an opposite solute-size dependence of SFE, i.e., a monotonic increase in SFE.[27] In the case, where the size of the solute is three times larger than the size of the solvent, the first peak of the solute-solvent pair distribution function calculated using the HNC approximation is significantly shifted towards shorter distances compared with the result using MD simulation.[27] RMDFT(D) and RMDFT(T) can improve these drawbacks in the HNC approximation for both the solute size dependence of SFE and the first peak position in the solute-solvent pair distribution function. However, in the case where the size of the solute is either larger than or different from that of the solvent, our results imply that the third-order bridge function might lead to the deterioration of the accuracy for both SFE and the solute-solvent pair distribution function.

In our previous study, we assessed the validity of the SFE functional that was derived from RMDFT(D) for an interaction-site model of water. The SFE functional with a reference HS fluid yielded results that were in good agreement with the experimental data for a set of neutral amino acid side-chain analogues, and 504 small organic molecules, although the values of SFE for hydrophilic solute molecules were slightly underestimated.[28,29] In these SFE calculations, the second-order direct correlation functions for bulk water and the density distribution functions of water around solute molecules,

which are needed as the input for the SFE calculation, were determined using a three-dimensional reference-interaction-site-model (3D-RISM) integral equation with a partially linearized HNC closure.[16] Using the 3D-RISM with the HNC-type closure is one of the reasons for the underestimation of the SFE values for hydrophilic solutes by the RMDFT functional. In future work, we will apply the RMDFT approach using the second- or third-order density functional Taylor series expansion to calculate both the direct correlations for bulk water and the solvent density distribution functions around solute molecules. We should then achieve further improvements in the theoretical prediction of SFE.

ACKNOWLEDGMENTS

This work was supported in part by a Grant-in-Aid for Scientific Research (KAKENHI) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

APPENDIX A: THE OZ RELATIONS FOR TRIPLET CORRELATION FUNCTIONS

The OZ relations for the triplet correlation functions (OZ3) are generated by using the functional differentiation of Eq. (38) and then applying the chain rule to the resulting relations.[33,52] The OZ3 equation has four alternative forms. One of them is given by

$$\frac{\delta n(\mathbf{r}_1|U)}{\delta[-\beta U(\mathbf{r}_2)]\delta[-\beta U(\mathbf{r}_3)]} = -\int d\mathbf{r}_4\, d\mathbf{r}_5\, d\mathbf{r}_6 \frac{\delta n(\mathbf{r}_1|U)}{\delta[-\beta U(\mathbf{r}_4)]} \frac{\delta[-\beta U(\mathbf{r}_4)]}{\delta n(\mathbf{r}_5|U)\delta n(\mathbf{r}_6|U)} \frac{\delta n(\mathbf{r}_6|U)}{\delta[-\beta U(\mathbf{r}_3)]} \frac{\delta n(\mathbf{r}_5|U)}{\delta[-\beta U(\mathbf{r}_2)]}, \quad \text{(A1)}$$

and the corresponding OZ3 relation is given as follows:

$$\begin{aligned} h^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = {} & C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) + h_{CA}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) \\ & + n_0 \int d\mathbf{r}_5 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_5|) C^{(3)}(\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_5) \\ & + n_0 \int d\mathbf{r}_6 h^{(2)}(|\mathbf{r}_2-\mathbf{r}_6|) C^{(3)}(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_6) \\ & + n_0 \int d\mathbf{r}_4 h^{(2)}(|\mathbf{r}_3-\mathbf{r}_4|) C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_4) \\ & + n_0^2 \int d\mathbf{r}_5\, d\mathbf{r}_6 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_5|) h^{(2)}(|\mathbf{r}_2-\mathbf{r}_6|) C^{(3)}(\mathbf{r}_3,\mathbf{r}_5,\mathbf{r}_6) \\ & + n_0^2 \int d\mathbf{r}_4\, d\mathbf{r}_5 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_5|) h^{(2)}(|\mathbf{r}_3-\mathbf{r}_4|) C^{(3)}(\mathbf{r}_2,\mathbf{r}_4,\mathbf{r}_6) \\ & + n_0^2 \int d\mathbf{r}_4\, d\mathbf{r}_6 h^{(2)}(|\mathbf{r}_2-\mathbf{r}_6|) h^{(2)}(|\mathbf{r}_3-\mathbf{r}_4|) C^{(3)}(\mathbf{r}_1,\mathbf{r}_4,\mathbf{r}_6) \\ & + n_0^3 \int d\mathbf{r}_4\, d\mathbf{r}_5\, d\mathbf{r}_6 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_5|) h^{(2)}(|\mathbf{r}_2-\mathbf{r}_6|) h^{(2)}(|\mathbf{r}_3-\mathbf{r}_4|) C^{(3)}(\mathbf{r}_4,\mathbf{r}_5,\mathbf{r}_6), \end{aligned} \quad \text{(A2)}$$

where we used the following relations:

$$\left.\frac{\delta n(\mathbf{r}_1|U)}{\delta[-\beta U(\mathbf{r}_2)]\delta[-\beta U(\mathbf{r}_3)]}\right|_{T,V,U=0} = n_0\delta(|\mathbf{r}_1-\mathbf{r}_3|)\delta(|\mathbf{r}_2-\mathbf{r}_3|)$$
$$+n_0^2\delta(|\mathbf{r}_2-\mathbf{r}_3|)h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|)$$
$$+n_0^2\delta(|\mathbf{r}_1-\mathbf{r}_2|)h^{(2)}(|\mathbf{r}_1-\mathbf{r}_3|)$$
$$+n_0^2\delta(|\mathbf{r}_1-\mathbf{r}_3|)h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|)$$
$$+n_0^3 g^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)-n_0^3 g^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|)$$
$$-n_0^3 h^{(2)}(|\mathbf{r}_1-\mathbf{r}_3|)-n_0^3 h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|)\,, \tag{A3}$$

$$g^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = 1+h^{(2)}(|\mathbf{r}_1-\mathbf{r}_2|)+h^{(2)}(|\mathbf{r}_1-\mathbf{r}_3|)+h^{(2)}(|\mathbf{r}_2-\mathbf{r}_3|)+h^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)\,, \tag{A4}$$

$$C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) \equiv -\beta\left.\frac{\delta^3 F^{ex}[n]}{\delta n(\mathbf{r}_1|U)\delta n(\mathbf{r}_2|U)\delta n(\mathbf{r}_3|U)}\right|_{T,V,U=0}$$
$$= \left.\frac{\delta^2[-\beta U_{IG}(\mathbf{r}_1)]}{\delta n_{IG}(\mathbf{r}_2|U_{IG})\delta n_{IG}(\mathbf{r}_3|U_{IG})}\right|_{T,V,U=0} - \left.\frac{\delta^2[-\beta U_{IG}(\mathbf{r}_1)]}{\delta n(\mathbf{r}_2|U)\delta n(\mathbf{r}_3|U)}\right|_{T,V,U=0}\,, \tag{A5}$$

$$\left.\frac{\delta^2[-\beta U_{IG}(\mathbf{r}_1)]}{\delta n_{IG}(\mathbf{r}_2|U_{IG})\delta n_{IG}(\mathbf{r}_3|U_{IG})}\right|_{T,V,U=0} = -\frac{1}{n_0^3}\left.\frac{\delta^2 n_{IG}(\mathbf{r}_1|U_{IG})}{\delta[-\beta U_{IG}(\mathbf{r}_2)]\delta[-\beta U_{IG}(\mathbf{r}_3)]}\right|_{T,V,U_{IG}=0}\,. \tag{A6}$$
$$= -\frac{1}{n_0}\delta(|\mathbf{r}_1-\mathbf{r}_3|)\delta(|\mathbf{r}_1-\mathbf{r}_2|)$$

APPENDIX B: A FACTORIZATION APPROXIMATION FOR $C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$

In order to understand the OZ3 equation given by Eq. (A2), a diagrammatical approach is more illustrative than the algebraic one. Thus, we define Lee's representation[33] as follows:

$h^{(2)}(\mathbf{r}_1,\mathbf{r}_2) =$

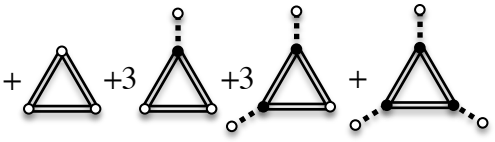
, (A7)

$C^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) =$ 

, (A8)

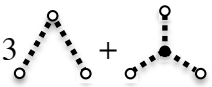
where the labeled point is denoted as an open circle.

The OZ3 equation of Eq. (A2) is expressed as

$h^{(3)} = h_{CA}^{(3)} +$ 

, (A9)

where the integral point, $n_0\int d\mathbf{r}_1$ is denoted as a solid circle and $h_{CK}^{(3)}$ defined by Eq. (41) is expressed as

$h_{CA}^{(3)} = 3$  .

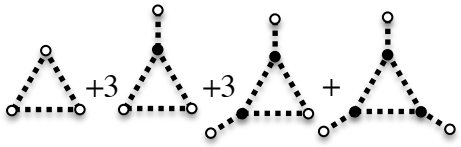
(A10)

If we apply the following factorization approximation to $C^{(3)}$ [41,42,47]

$C_{CK}^{(3)} =$  ,

(A11)

we obtain $\Delta h_{CK}^{(3)}$, defined by Eq. (43), as follows:

$\Delta h_{CK}^{(3)} =$  .

(A12)

If we apply the Kirkwood superposition (KS) approximation to $g^{(3)}\left(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3\right)$, we obtain

$$g_{KS}^{(3)}\left(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3\right) = g^{(2)}\left(\mathbf{r}_1,\mathbf{r}_2\right)g^{(2)}\left(\mathbf{r}_1,\mathbf{r}_3\right)g^{(2)}\left(\mathbf{r}_2,\mathbf{r}_3\right). \tag{A13}$$

The triplet correlation function, $h_{KS}^{(3)}\left(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3\right)$, for the KS approximation is expressed as

$h_{KS}^{(3)} = 3$  .

(A14)

Equation (A14) corresponds to the sum of $C_{CK}^{(3)}$ and the first terms of $h_{CK}^{(3)}$.

As pointed out by Iyetomi and Ichimaru,[41,42,47] the additional terms given by Eq. (A12) can be regarded as a sum of $C_{CK}^{(3)}$ and all the vertex corrections of $C_{CK}^{(3)}$. In the case of bulk liquids, the third-order bridge function, where $C_{CK}^{(3)}$ is approximately used as $C_{ex}^{(3)}$, that is, Eq. (55), is given by[47]

$B_{CK}^{(3)} = \frac{1}{2}$ 

(A15)

Equation (A15) is the simplest elementary bridge diagram with a $h^{(2)}$-bond. This observation suggests that the factorization approximation for $C^{(3)}$ given by Eq. (A11) is valid for at least homogeneous liquids.

APPENDIX C: COMPUTATIONAL DETAILS FOR SOLVING THE INTEGRAL EQUATIONS

All the integral equations were solved using $r_{\max} = 50d_v$ as the maximum radial distance, where d_v is the Lennard-Jones diameter for the bulk/solvent system. In total, 4096 grid points were used to discretize the integral equations. In the present study, the HS diameter for the reference system was chosen as $d_{HS} = 1.013d_v$, so that the structure factor, $S(k)$, at small-k values became close to the MD result of a LJ liquid near the triple point.[50]

APPENDIX D: AN EFFICIENT NEUMERICAL INTEGRATION OF $B_{CK}^{(3)}$

As shown in Eq. (A15), since the third-order bridge function, $B_{CK}^{(3)}$, is the bridge diagram, we need an efficient numerical integration to evaluate $B_{CK}^{(3)}$. The multiple integrals included in $B_{CK}^{(3)}$ can be reduced to simple double integrals by Legendre polynomials expansion.[41,53] This technique was proposed by Barker and Monaghan in order to perform the numerical calculation of the fourth virial coefficient.[53] $h^{(2)}(r)$ can be expanded in terms of Legendre polynomials:

$$h^{(2)}\left(\left|\mathbf{r}_1 - \mathbf{r}_2\right|\right) = \sum_{l=0}^{\infty} A_l(r_1, r_2) P_l(\cos\theta_{12}), \tag{A16}$$

where θ_{12} is the angle between vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. The expansion coefficients $A_l(r_1, r_2)$ are given by

$$A_l(r_1, r_2) = \frac{2l+1}{2} \int_{-1}^{1} h^{(2)}\left(\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta_{12}}\right) P_l(\cos\theta_{12}) d(\cos\theta_{12}) . \tag{A17}$$

If we substitute the expansion (Eq. (A16)) into Eq. (55) and then use the orthogonal property and additional theorem for the Legendre polynomials, the angle-dependent part of $B_{CK}^{(3)}$ can be simplified. As a result, we obtain the tractable expression that includes only double integrals with respect to r_1 and r_2:

$$B_{CK}^{(3)}(r) = 8\pi^2 \sum_{l=0}^{\infty} \frac{1}{(2l+1)^2} \int_0^{\infty} dr_1 \int_0^{\infty} dr_2 r_1^2 r_2^2 A_l(r, r_1) A_l(r, r_2) A_l(r_1, r_2) \Delta n\left(r_1 \middle| U_{PR}\right) \Delta n\left(r_2 \middle| U_{PR}\right) . \tag{A18}$$

In the present study, we took into account the polynomial series expansion in Eq. (A18) up to $l_{\max} = 25$ to get satisfactory convergence. The maximum radial distances for the double integrals with respect to r_1 and r_2 in Eq. (A18) are $5d_v$ and $10d_v$ for the bulk/pure solvent and solute-solvent systems, respectively. The numerical integrations in Eq. (A17) were performed using the Gauss-Legendre quadrature of 25 points.

REFERENCES

[1] A. Ben-Naim, Current Opinion in Structural Biology **4**, 264 (1993).
[2] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1989).
[3] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 2001).
[4] W.L. Jorgensen and C. Ravimohan, The Journal of Chemical Physics **83**, 3050 (1985).
[5] T.P. Straatsma and H.J.C. Berendsen, The Journal of Chemical Physics **89**, 5876 (1988).
[6] M.R. Shirts, J.W. Pitera, W.C. Swope, and V.S. Pande, J Chem Phys **119**, 5740 (2003).
[7] M.R. Shirts and V.S. Pande, J Chem Phys **122**, 134508 (2005).
[8] N. Matubayasi and M. Nakahara, The Journal of Chemical Physics **113**, 6070 (2000).
[9] N. Matubayasi and M. Nakahara, J Chem Phys **117**, 3605 (2002).
[10] N. Matubayasi and M. Nakahara, The Journal of Chemical Physics **119**, 9686 (2003).
[11] H.C. Andersen and D. Chandler, The Journal of Chemical Physics **57**, 1918 (1972).
[12] D. Chandler and H.C. Andersen, The Journal of Chemical Physics **57**, 1930 (1972).
[13] F. Hirata, J Chem Phys **77**, 509 (1982).
[14] B.M. Pettitt, J Chem Phys **77**, 1451 (1982).
[15] S.J. Singer and D. Chandler, Molecular Physics **55**, 621 (1985).
[16] A. Kovalenko and F. Hirata, The Journal of Chemical Physics **110**, 10095 (1999).
[17] G.N. Chuev, M.V. Fedorov, and J. Crain, Chemical Physics Letters **448**, 198 (2007).
[18] E.L. Ratkova, G.N. Chuev, V.P. Sergiievskyi, and M.V. Fedorov, J Phys Chem B **114**, 12068 (2010).
[19] J.-F. Truchon, B.M. Pettitt, and P. Labute, J. Chem. Theory Comput. **10**, 934 (2014).
[20] V.P. Sergiievskyi, G. Jeanmairet, M. Levesque, and D. Borgis, J. Phys. Chem. Lett. **5**, 1935 (2014).
[21] Y. Liu, S. Zhao, and J. Wu, J. Chem. Theory Comput. **9**, 1896 (2013).
[22] F. Hirata and P.J. Rossky, Chemical Physics Letters **83**, 329 (1981).
[23] J.-P. Hansen and I.R. McDonald, *Theory of Simple Liquids; 2nd Ed.* (Academic Press, London, 1986).
[24] S. Ten-no, The Journal of Chemical Physics **115**, 3724 (2001).
[25] K. Sato, H. Chuman, and S. Ten-no, J Phys Chem B **109**, 17290 (2005).
[26] A. Kovalenko and F. Hirata, J Chem Phys **113**, 2793 (2000).
[27] T. Miyata and J. Thapa, Chemical Physics Letters **604**, 122 (2014).
[28] T. Sumi, A. Mitsutake, and Y. Maruyama, J. Comput. Chem. **36**, 1359 (2015).
[29] T. Sumi, A. Mitsutake, and Y. Maruyama, J. Comput. Chem. **36**, 2009 (2015).
[30] Y. Rosenfeld and N.W. Ashcroft, Physical Review a (General Physics) **20**, 1208 (1979).
[31] D. Henderson, *Fundamentals of Inhomogeneous Fluids* (Marcel Dekker, New York, 1992).
[32] R.G. Parr and Y. Weitao, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1994).
[33] L.L. Lee, The Journal of Chemical Physics **60**, 1197 (1974).
[34] W.A. Curtin, J Chem Phys **88**, 7050 (1988).
[35] Y. Rosenfeld, Phys. Rev. Lett. **63**, 980 (1989).
[36] R. Roth, R. Evans, A. Lang, and G. Kahl, J. Phys.: Condens. Matter **14**, 12063 (2002).
[37] Y.-X. Yu and J. Wu, J Chem Phys **117**, 10156 (2002).
[38] T. Sumi and H. Sekino, J. Phys. Soc. Jpn. **77**, 034605 (2008).
[39] J.K. Percus, Phys. Rev. Lett. **8**, 462 (1962).
[40] Y. Rosenfeld, J Chem Phys **98**, 8126 (1993).
[41] H. Iyetomi and S. Ichimaru, Physical Review a (General Physics) **25**, 2434 (1982).
[42] H. Iyetomi and S. Ichimaru, Physical Review a (General Physics) **27**, 3241 (1983).
[43] J. Barrat, J. Hansen, and G. Pastore, Phys. Rev. Lett. **58**, 2075 (1987).
[44] J.L. Barrat, J.P. Hansen, and G. Pastore, Molecular Physics **63**, 747 (1988).
[45] L.L. Lee, J Chem Phys **135**, 204706 (2011).

[46] S. Ichimaru, Phys. Rev., A **2**, 494 (1970).
[47] S. Ichimaru, Reviews of Modern Physics **54**, 1017 (1982).
[48] Y. Rosenfeld, D. Levesque, and J.J. Weis, J Chem Phys **92**, 6818 (1990).
[49] N.F. Carnahan and K.E. Starling, The Journal of Chemical Physics **51**, 635 (1969).
[50] L. Reatto, D. Levesque, and J. Weis, Phys. Rev., A **33**, 3451 (1986).
[51] K. Koga, Phys Chem Chem Phys **13**, 19749 (2011).
[52] J.L. Lebowitz and J.K. Percus, Journal of Mathematical Physics **4**, 116 (1963).
[53] J.A. Barker and J.J. Monaghan, J Chem Phys **36**, 2564 (1962).