The BGY3dM model for the approximation of solvent densities

Michael Griebel[a] and Lukas Jager[b]
Institut für Numerische Simulation, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany

(Received 11 June 2008; accepted 9 September 2008; published online 6 November 2008)

We present a new approach for the approximation of solvent densities around solutes of arbitrary shape. Our model represents a three-dimensional (3d) Born–Green–Yvon (BGY) equation for an arbitrary solute immersed into a molecular (M) solvent, the BGY3dM model. It comprises the famous Kirkwood approximation as closure relation. The molecules of the solvent are modeled as rigid bodies by taking the limit of an infinite restoring force for the intramolecular interactions. Furthermore, short-range potentials as well as the long-range Coulomb interaction are taken into account. The resulting integro-differential equations are efficiently solved by a Picard iteration and a solution of the linearized equations using Fourier transformations. We compare the results obtained from the presented BGY3dM method with results obtained by extensive molecular dynamics simulations for a HCl-like model solvent. Furthermore, we apply the method to carbon disulfide as solvent. The overall performance of the method is promising. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2991296]

I. INTRODUCTION

The microscopic simulation of molecules such as proteins in solution is a challenging task. An explicit simulation of the entire solute-solvent system is often unfeasible due to the high number of degrees of freedom needed to adequately simulate the solvent effects.[1] Therefore, implicit solvent models have been developed which take the influence of the solvent into consideration without explicitly introducing new degrees of freedom to the system. Most of these implicit solvent models approximate the solvent effects by a continuum model which clearly neglects important local properties of the solvent.[2,3] Hence, the development of new implicit solvent models which approximate the solvent effects more accurately is a key topic of current research.

The simulation of a solute-solvent system can be formally simplified by introducing the potential of mean force (PMF). However, the PMF can only be computed efficiently if the mean solvent density around the solute is known.[2] Promising developments for the fast approximation of the mean solvent density were made by the application of the liquid state integral equation theories. Several authors developed methods based on these theories which allow to compute solvent densities around solutes of arbitrary shape for simple monoatomic[4,5] as well as for molecular solvents.[6–14] It stands out that practically all such methods found in the literature are based on the Ornstein–Zernike equation together with the so-called reference interaction site model[15] (RISM) and mostly employ the hypernetted chain (HNC) closure. However, these methods do not lead to a satisfactory accuracy in all situations.[6–14] Additionally, the computational effort involved still makes a repeated evaluation during an extensive solute-solvent simulation unfeasible.

To our knowledge, methods based directly on the Yvon–Born–Green (YBG) hierarchy have never been considered for the computation of solvent densities in solute-solvent systems. Methods that are related to the computation of pair distribution functions of pure molecular fluids were developed in the field of polymeric fluids only. To this end, Eu and Gan,[16] Taylor and Lipson,[17] and Attard[18] have derived equations based on the YBG hierarchy that have been quite successfully applied to several polymer models.[19–27] In these models, a polymer chain consists of either hard or soft spheres with rigid or flexible bonds. But neither chains with different types of particles nor more complex interaction potentials as, e.g., the Coulomb potential have been considered.

In this paper, we present a new approach derived directly from the YBG hierarchy which employs the Kirkwood superposition approximation.[28] More specifically, it represents a three-dimensional (3d) Born–Green–Yvon (BGY) equation for an arbitrary solute immersed into a molecular (M) solvent and is hence called the BGY3dM model. We will investigate its properties and benefits for the computation of solvent density distributions around a solute of arbitrary shape. With our new model we are able to treat molecular solvents that interact with the solute by any short-range potential and the long-range Coulomb potential. The solvent molecules are modeled as rigid bodies. A comparison between the BGY3dM method and results obtained by extensive molecular dynamics (MD) simulations demonstrate the quite good performance of the new method.

The remainder of this article is organized as follows: First, we derive the BGY3dM model in Sec. II before we present some computational details in Sec. III. Then, results from the BGY3dM model are compared with results from MD simulations for a HCl-like model solvent and the method is applied to carbon disulfide as solvent in Sec. IV. Finally, we give a short summary and outlook of the presented method in Sec. V.

[a]Electronic mail: griebel@ins.uni-bonn.de.
[b]Electronic mail: jager@ins.uni-bonn.de.

II. THE BGY3DM MODEL

In the following, we present our BGY3dM model for the approximation of solvent densities around an arbitrary solute. To this end, we first derive the site-site BGY3dM (SS-BGY3dM) equations for the computation of pair distribution functions of the pure solvent, since these pair distribution functions are needed as input for the BGY3dM equations, which will be described afterwards. Both the SS-BGY3dM and the BGY3dM models represent equations from two distinct YBG hierarchies, i.e., that for pure molecular fluids and that for a solute immersed into a molecular fluid, respectively. These hierarchies are obtained by integrating the corresponding Liouville equations similar to the monoatomic case.[29] The application of approximations for the distribution functions then yields a closed set of equations, which can be solved numerically.

A. The site-site BGY3dM equations

We begin our derivation with the BGY equation for a simple mixture of s different particle species as it is known from the literature.[30] It reads as

$$\nabla_{\mathbf{x}_1^\alpha} g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) = \beta \mathbf{F}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) + \beta \sum_{\eta=1}^{s} \bar{\rho}_\eta \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta) \times g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma,\mathbf{x}_3^\eta) d\mathbf{x}_3^\eta \tag{1}$$

for any pair $\alpha,\gamma=1,\ldots,s$. Here, $g^{(2)}_{\alpha\gamma}$ and $g^{(3)}_{\alpha\gamma\eta}$ are the pair and triplet distribution functions for particles of species α, γ, and η, respectively, $\mathbf{F}_{\alpha\gamma}=-\nabla_{\mathbf{x}_1^\alpha} v_{\alpha\gamma}$ denotes the force between particles of species α and γ, and $\bar{\rho}_\alpha$ is the number density of particle species α. For clarity the position vectors have a superscript also indicating their respective particle species, e.g., $\mathbf{x}_1^\alpha \in \mathbb{R}^3$.

We now consider molecular solvents. Therefore, we set $\bar{\rho}_S := \bar{\rho}_1 = \bar{\rho}_2 = \ldots = \bar{\rho}_s$. Then, we build molecules which contain exactly one particle from every not necessarily different species. By this, we model a molecular fluid with number density $\bar{\rho}_S$, where the molecules consist of s particles. When we now derive the BGY equation for this molecular fluid, we have to distinguish between intramolecular and intermolecular interactions. The intramolecular forces are indicated by the superscript i, i.e., we write $\mathbf{F}^i_{\alpha\gamma}$ and $v^i_{\alpha\gamma}$ for the forces and the pair potential, respectively. Similar to the forces, we now have different types of distribution functions, which depend on how many of the particles belong to the same molecule. We indicate this dependency by the indices of the corresponding position vectors, i.e., $\mathbf{x}_j^\alpha$ denotes the position of particle α of molecule j. Hence, $g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma)$ denotes the intramolecular pair distribution function between particle α and γ of the same molecule whereas $g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)$ denotes the intermolecular pair distribution function between particles α and γ of different molecules. In this notation we can write the BGY equation for the intermolecular pair distribution functions $(n=2)$ of a molecular fluid as

$$\begin{aligned}\nabla_{\mathbf{x}_1^\alpha} g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) &= \beta \mathbf{F}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) \\ &+ \beta \sum_{\eta=1}^{s} \bar{\rho}_S \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta) g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma,\mathbf{x}_3^\eta) d\mathbf{x}_3^\eta \\ &+ \beta \sum_{\eta=1,\eta\neq\alpha}^{s} \int_\Omega \mathbf{F}^i_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta) d\mathbf{x}_1^\eta \\ &+ \beta \sum_{\eta=1,\eta\neq\gamma}^{s} \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma,\mathbf{x}_2^\eta) d\mathbf{x}_2^\eta, \end{aligned} \tag{2}$$

and the BGY equation for the intramolecular pair distribution functions as

$$\begin{aligned}\nabla_{\mathbf{x}_1^\alpha} g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma) &= \beta \mathbf{F}^i_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma) \\ &+ \beta \sum_{\eta=1}^{s} \bar{\rho}_S \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma,\mathbf{x}_2^\eta) d\mathbf{x}_2^\eta \\ &+ \beta \sum_{\eta=1,\eta\neq\alpha,\eta\neq\gamma}^{s} \int_\Omega \mathbf{F}^i_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma,\mathbf{x}_1^\eta) d\mathbf{x}_1^\eta. \end{aligned} \tag{3}$$

As for a simple mixture, the equations are obtained by integrating the Liouville equation over $N-2$ particles, i.e., we choose the second equation of the YBG hierarchy for molecular fluids. To this end, one has to take into account that particles can now belong to different molecules. This additional distinguishability results in the last two lines of Eq. (2) and the last line of Eq. (3) which represent the intramolecular coupling within the molecules. The first lines are identical to those of a simple mixture as in Eq. (1). Recall that all particle species in the molecular fluid have the same number density $\bar{\rho}_S$.

In order to facilitate the numerical solution of the final equations, the molecules are modeled as rigid bodies. To this end, we introduce a harmonic potential as intramolecular interaction

$$v^i(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa) = \kappa(r_1^{\alpha\gamma} - r_0^{\alpha\gamma})^2, \quad \forall\ \alpha \neq \gamma, \tag{4}$$

with $r_1^{\alpha\gamma}=|\mathbf{x}_1^\alpha-\mathbf{x}_1^\gamma|$. Here, $r_0^{\alpha\gamma}$ denotes the desired intramolecular distance between particles of species α and γ. The constant κ defines the strength of the potential. The $s(s-1)/2$ different distances $r_0^{\alpha\gamma}$ completely specify the configuration of the molecule. The potential (4) does not yet lead to fixed distances within the molecule, but allows fluctuations around the desired distances $r_0^{\alpha\gamma}$. Hence, we investigate the limit case where the constant κ goes to infinity, i.e., we consider $\lim_{\kappa\to\infty} v^i(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa)$. Here, κ determines the strength of the force that constrains two particles to their desired distance. We examine Eq. (3), which determines the intramolecular pair distribution functions, and assume that, in this limit, the solution of Eq. (3) is strongly dominated by the first term of the right-hand side, and that all integral terms can be neglected, i.e., for $\kappa\to\infty$ we have

$$\nabla_{\mathbf{x}_1^\alpha} g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa) = \beta \mathbf{F}^i_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa). \tag{5}$$

The dependence on κ is explicitly written in the arguments in order to distinguish between $g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa)$ and the limit version of $g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma)$ which will not depend on this parameter. The solution of Eq. (5) is

$$\begin{aligned} g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa) &= \sqrt{\frac{4\beta\kappa}{\pi}} \frac{1}{4\pi(r_0^{\alpha\gamma})^2} \exp(-\beta v^i(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma;\kappa)) \\ &= \sqrt{\frac{4\beta\kappa}{\pi}} \frac{1}{4\pi(r_0^{\alpha\gamma})^2} \exp(-\beta\kappa(r_1^{\alpha\gamma}-r_0^{\alpha\gamma})^2). \end{aligned} \tag{6}$$

The factor $\sqrt{(4\beta\kappa/\pi)}(1/4\pi(r_0^{\alpha\gamma})^2)$ is chosen such that the intramolecular pair distribution function obeys the correct normalization condition

$$\lim_{\kappa\to\infty} \int_\Omega g^{(2)}_{\alpha\gamma}(\mathbf{r}_1^{\alpha\gamma};\kappa) d\mathbf{r}_1^{\alpha\gamma} = 1, \tag{7}$$

with $\mathbf{r}_1^{\alpha\gamma}=\mathbf{x}_1^\alpha-\mathbf{x}_1^\gamma$. With this choice, we find for the convolution with an arbitrary function f

$$\begin{aligned} &\int_\Omega f(\mathbf{r}_1^{\alpha\gamma}-\mathbf{r}') g^{(2)}_{\alpha\gamma}(\mathbf{r}_1^{\alpha\gamma}) d\mathbf{r}_1^{\alpha\gamma} \\ &= \lim_{\kappa\to\infty} \int_\Omega f(\mathbf{r}_1^{\alpha\gamma}-\mathbf{r}') g^{(2)}_{\alpha\gamma}(\mathbf{r}_1^{\alpha\gamma};\kappa) d\mathbf{r}_1^{\alpha\gamma} \\ &= \lim_{\kappa\to\infty} \int_\Omega f(\mathbf{r}_1^{\alpha\gamma}-\mathbf{r}') \sqrt{\frac{4\beta\kappa}{\pi}} \frac{\exp(-\beta\kappa(r_1^{\alpha\gamma}-r_0^{\alpha\gamma})^2)}{4\pi(r_0^{\alpha\gamma})^2} d\mathbf{r}_1^{\alpha\gamma} \\ &= \int_\Omega f(\mathbf{r}_1^{\alpha\gamma}-\mathbf{r}') \frac{\delta(r_1^{\alpha\gamma}-r_0^{\alpha\gamma})}{4\pi(r_0^{\alpha\gamma})^2} d\mathbf{r}_1^{\alpha\gamma}. \end{aligned} \tag{8}$$

Equation (8) represents the definition of the delta distribution as the limit of a Dirac sequence. The result is intuitive, since the two particles $\mathbf{x}_1^\alpha$ and $\mathbf{x}_1^\gamma$ have to remain exactly at a distance of $r_0^{\alpha\gamma}$ if the restoring force is infinite. The factor $4\pi(r_0^{\alpha\gamma})^2$ represents the surface of the sphere with radius $r_0^{\alpha\gamma}$ and ensures the correct normalization. Consequently, we know all *intra*molecular pair distribution functions and set

$$g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\gamma) = \frac{\delta(r_1^{\alpha\gamma}-r_0^{\alpha\gamma})}{4\pi(r_0^{\alpha\gamma})^2}, \quad \forall\ \alpha \neq \gamma. \tag{9}$$

Now, to solve Eq. (2) for the *inter*molecular pair distribution functions, a closure relation is needed. For this, we approximate the triplet distribution functions by products of functions that depend only on two of the three particle coordinates. In the case where all three particles belong to different molecules we insert the famous Kirkwood approximation[28]

$$g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma,\mathbf{x}_3^\eta) \approx g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta) g^{(2)}_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_3^\eta). \tag{10}$$

In the case where two of the particles belong to the same molecule, the Kirkwood approximation is not satisfactory.[18] Hence, we employ the so-called normalized site-site superposition approximation (NSSA) of Taylor and Lipson[17]

$$g^{(3)}_{\alpha\gamma\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma,\mathbf{x}_2^\eta) \approx g^{(2)}_{\gamma\eta}(x_2^\gamma,\mathbf{x}_2^\eta) \tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) \tilde{g}^{(2)}_{\alpha\eta,\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta), \tag{11}$$

with

$$\begin{aligned} \tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) &= \frac{g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)}{n^\eta_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)}, \\ \tilde{g}^{(2)}_{\alpha\eta,\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) &= \frac{g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta)}{n^\gamma_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta)} \end{aligned} \tag{12}$$

and the normalization functions

$$\begin{aligned} \eta^\eta_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) &= \int_\Omega g^{(2)}_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_2^\eta) g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) d\mathbf{x}_2^\eta, \\ n^\gamma_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) &= \int_\Omega g^{(2)}_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_2^\eta) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) d\mathbf{x}_2^\gamma. \end{aligned} \tag{13}$$

Applying the Kirkwood approximation and the NSSA approximation for the intermolecular and intramolecular terms of Eq. (2) finally gives

$$\begin{aligned} \nabla_{\mathbf{x}_1^\alpha} g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) &= \beta \mathbf{F}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) + \beta g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) \sum_{\eta=1}^{s} \bar{\rho}_S \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta) g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta) g^{(2)}_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_3^\eta) d\mathbf{x}_3^\eta \\ &+ \beta \tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) \sum_{\eta=1,\eta\neq\alpha}^{s} \int_\Omega \mathbf{F}^i_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) \tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta) d\mathbf{x}_1^\eta \\ &+ \beta \tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma) \sum_{\eta=1,\eta\neq\gamma}^{s} \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) \tilde{g}^{(2)}_{\alpha\eta,\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g^{(2)}_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_2^\eta) d\mathbf{x}_2^\eta. \end{aligned} \tag{14}$$

In the last line of Eq. (14) we can now substitute the intramolecular pair distribution function according to Eq. (9). The term which includes the intramolecular force $\mathbf{F}^i_{\alpha\eta}$, however, requires a more detailed examination of the limit $\kappa\to\infty$. If we insert the definition of the intramolecular potential (4) and the intramolecular pair distribution from Eq. (6), we have

$$\lim_{\kappa\to\infty}\sum_{\eta=1,\eta\neq\alpha}^{s}\beta\tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)\int_\Omega \mathbf{F}^i_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta;\kappa)g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta;\kappa)\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta)d\mathbf{x}_1^\eta$$

$$=\lim_{\kappa\to\infty}\sum_{\eta=1,\eta\neq\alpha}^{s}\beta\tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)\int_\Omega(-\nabla_{\mathbf{x}_1^\alpha}\kappa(r_1^{\alpha\eta}-r_0^{\alpha\eta})^2)\sqrt{\frac{4\beta\kappa}{\pi}}\frac{\exp(-\beta\kappa(r_1^{\alpha\eta}-r_0^{\alpha\eta})^2)}{4\pi(r_0^{\alpha\eta})^2}\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta)d\mathbf{x}_1^\eta$$

$$=\lim_{\kappa\to\infty}\sum_{\eta=1,\eta\neq\alpha}^{s}\tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)\int_\Omega\left(\nabla_{\mathbf{x}_1^\alpha}\sqrt{\frac{4\beta\kappa}{\pi}}\frac{\exp(-\beta\kappa(r_1^{\alpha\eta}-r_0^{\alpha\eta})^2)}{4\pi(r_0^{\alpha\eta})^2}\right)\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta)d\mathbf{x}_1^\eta$$

$$=\lim_{\kappa\to\infty}\sum_{\eta=1,\eta\neq\alpha}^{s}\tilde{g}^{(2)}_{\alpha\gamma,\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)\int_\Omega\sqrt{\frac{4\beta\kappa}{\pi}}\frac{\exp(-\beta\kappa(r_1^{\alpha\eta}-r_0^{\alpha\eta})^2)}{4\pi(r_0^{\alpha\eta})^2}(\nabla_{\mathbf{x}_1^\eta}\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta))d\mathbf{x}_1^\eta$$

$$=\sum_{\eta=1,\eta\neq\alpha}^{s}\frac{g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)}{n^\eta_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)}\nabla_{\mathbf{x}_1^\alpha}\int_\Omega\frac{\delta(r_1^{\alpha\eta}-r_0^{\alpha\eta})}{4\pi(r_0^{\alpha\eta})^2}\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta)d\mathbf{x}_1^\eta\approx g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)\sum_{\eta=1,\eta\neq\alpha}^{s}\nabla_{\mathbf{x}_1^\alpha}\ln\left(\int_\Omega\frac{\delta(r_1^{\alpha\eta}-r_0^{\alpha\eta})}{4\pi(r_0^{\alpha\eta})^2}\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta)d\mathbf{x}_1^\eta\right). \tag{15}$$

In the last line of Eq. (15) we approximated $g^{(2)}_{\gamma\eta}$ in the definition of $n^\eta_{\alpha\gamma}$ by its normalized form $\tilde{g}^{(2)}_{\gamma\eta,\alpha}$, such that we can write the fraction as the gradient of a logarithm.

A further simplification of Eq. (14) is possible by introducing a product ansatz for the solution

$$g^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)=g^0_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)\exp(-u^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)), \tag{16}$$

with

$$g^0_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)=\exp(-\beta v_{\alpha\gamma}(\mathbf{x}_1^\alpha-\mathbf{x}_2^\gamma)), \tag{17}$$

where $v_{\alpha\gamma}(\mathbf{x}_1^\alpha-\mathbf{x}_2^\gamma)$ denotes the potential between sites α and γ of molecules one and two, respectively. Here, $u^{(2)}_{\alpha\gamma}$ is the new unknown function. This approach together with the application of the divergence to both sides of the equation leads after some calculation to the following equation for the intermolecular site-site pair distribution function:

$$\Delta_{\mathbf{x}_1^\alpha}u^{(2)}_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)=\beta\sum_{\eta=1}^{s}\bar{\rho}_S\nabla_{\mathbf{x}_1^\alpha}\cdot\int_\Omega\mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta)g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_3^\eta)g^{(2)}_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_3^\eta)d\mathbf{x}_3^\eta$$

$$-\sum_{\eta=1,\eta\neq\alpha}^{s}\Delta_{\mathbf{x}_1^\alpha}\ln\left(\int_\Omega\omega_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta)\tilde{g}^{(2)}_{\gamma\eta,\alpha}(\mathbf{x}_2^\gamma,\mathbf{x}_1^\eta)d\mathbf{x}_1^\eta\right)$$

$$-\beta\sum_{\eta=1,\eta\neq\gamma}^{s}\nabla_{\mathbf{x}_1^\alpha}\cdot\frac{\int_\Omega\mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta)\tilde{g}^{(2)}_{\alpha\eta,\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta)\omega_{\gamma\eta}(\mathbf{x}_2^\gamma,\mathbf{x}_2^\eta)d\mathbf{x}_2^\eta}{n^\eta_{\alpha\gamma}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\gamma)}. \tag{18}$$

Here, we introduced the notation

$$\omega_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta):=g^{(2)}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta)=\frac{\delta(r_1^{\alpha\eta}-r_0^{\alpha\eta})}{4\pi(r_0^{\alpha\eta})^2} \tag{19}$$

for the intramolecular pair distribution functions. Note that we switch here to the ω notation to agree with the literature. The dot $(\cdot)$ in the right hand side of Eq. (18) indicates the scalar product of two vectors in $\mathbb{R}^3$. We call Eq. (18) for any pair $\alpha,\ \gamma=1,\ldots,s$ with $\alpha\neq\gamma$ the site-site BGY3dM (SS-BGY3dM) equations.

B. The BGY3dM equations

Next, we consider a single solute molecule immersed into the solvent. We wish to compute the mean density of the solvent around a fixed configuration of the solute. If we assume that the solute is described by N_M particles with fixed configuration $\mathbf{x}^M:=(\mathbf{x}_1^M,\ldots,\mathbf{x}_{N_M}^M)$ the mean solvent density ρ^S_α of particle type α can be expressed by a conditional (N_M+1)-particle distribution

$$\rho_\alpha^S(\mathbf{x}_1^\alpha) := \bar{\rho}_S g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha|\mathbf{x}^M) = \bar{\rho}_S \frac{g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M)}{g^{(N_M)}(\mathbf{x}^M)}. \tag{20}$$

In analogy to the BGY equation for the site-site pair distribution functions of a molecular fluid we can derive the BGY equation for the (N_M+1)-particle distribution function where N_M particles belong to the solute,

$$\begin{aligned}\nabla_{\mathbf{x}_1^\alpha} g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M) &= \beta \mathbf{F}_\alpha(\mathbf{x}_1^\alpha,\mathbf{x}^M) g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M) \\ &+ \beta \sum_{\eta=1}^{s} \bar{\rho}_S \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g_{\alpha\eta}^{(N_M+2)}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta,\mathbf{x}^M) d\mathbf{x}_2^\eta \\ &+ \beta \sum_{\eta=1,\eta\neq\alpha}^{s} \int_\Omega \mathbf{F}_{\alpha\eta}^i(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) g_{\alpha\eta}^{(N_M+2)}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta,\mathbf{x}^M) d\mathbf{x}_1^\eta.\end{aligned} \tag{21}$$

To this end, $\mathbf{F}_\alpha(\mathbf{x}_1^\alpha,\mathbf{x}^M)$ is the total force exerted on the solvent particle $\mathbf{x}_1^\alpha$ due to the solute. Next, we employ the so-called n-level Kirkwood closure relations[31] and approximate the intermolecular and intramolecular (N_M+2)-particle distribution functions by

$$\begin{aligned} g_{\alpha\eta}^{(N_M+2)}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta,\mathbf{x}^M) &\approx \frac{g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M) g_{\alpha\eta}^{(2)}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g_\eta^{(N_M+1)}(\mathbf{x}_2^\eta,\mathbf{x}^M)}{g^{(N_M)}(\mathbf{x}^M)}, \\ g_{\alpha\eta}^{(N_M+2)}(\mathbf{x}_\eta^\alpha,\mathbf{x}_\eta^\eta,\mathbf{x}^M) &\approx \frac{\tilde{g}_{\alpha;\eta}^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M) \omega_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) \tilde{g}_{\eta;\alpha}^{(N_M+1)}(\mathbf{x}_1^\eta,\mathbf{x}^M)}{g^{(N_M)}(\mathbf{x}^M)},\end{aligned} \tag{22}$$

with

$$\begin{aligned}\tilde{g}_{\alpha;\eta}^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M) &:= \frac{g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha,\mathbf{x}^M)}{n_\alpha^\eta(\mathbf{x}_1^\alpha,\mathbf{x}^M)}, \\ \tilde{g}_{\eta;\alpha}^{(N_M+1)}(\mathbf{x}_1^\eta,\mathbf{x}^M) &:= \frac{g_\eta^{(N_M+1)}(\mathbf{x}_1^\eta;\mathbf{x}^M)}{n_\eta^\alpha(\mathbf{x}_1^\eta,\mathbf{x}^M)}\end{aligned} \tag{23}$$

and

$$\begin{aligned} n_\alpha^\eta(\mathbf{x}_1^\alpha) &:= \int_\Omega \omega_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) \tilde{g}_{\eta;\alpha}^{(N_M+1)}(\mathbf{x}_1^\eta|\mathbf{x}^M) d\mathbf{x}_1^\eta, \\ n_\eta^\alpha(\mathbf{x}_1^\eta) &:= \int_\Omega \omega_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha|\mathbf{x}^M) d\mathbf{x}_1^\alpha.\end{aligned} \tag{24}$$

For the approximation of the intramolecular distribution function we additionally simplified the NSSA approximation (22) of Taylor and Lipson[17] by replacing $g_\eta^{(N_M+1)}$ by its normalized form $\tilde{g}_{\eta;\alpha}^{(N_M+1)}$ in Eq. (24). Now, we can insert the above approximations into Eq. (21) and consider the modeling of the rigid bonds of the molecules in the same way as for the SS-BGY3dM equations. Then, we divide the whole equation by $g^{(N_M)}(\mathbf{x}^M)$ and employ the product approach,

$$g_\alpha(\mathbf{x}_1^\alpha) := g_\alpha^{(N_M+1)}(\mathbf{x}_1^\alpha|\mathbf{x}^M) = g_\alpha^0(\mathbf{x}_1^\alpha;\mathbf{x}^M)\exp(-u_\alpha(\mathbf{x}_1^\alpha)), \tag{25}$$

with

$$g_\alpha^0(\mathbf{x}_1^\alpha;\mathbf{x}^M) = \exp(-\beta v_\alpha(\mathbf{x}_1^\alpha - \mathbf{x}^M)), \tag{26}$$

where $v_\alpha(\mathbf{x}_1^\alpha - \mathbf{x}^M)$ describes the total potential between solvent particle α and the solute. Finally, we apply the divergence and obtain the equation

$$\begin{aligned}\Delta_{\mathbf{x}_1^\alpha} u_\alpha(\mathbf{x}_1^\alpha) = &-\beta \sum_{\eta=1}^{s} \bar{\rho}_S \nabla_{\mathbf{x}_1^\alpha} \\ &\cdot \int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g_{\alpha\eta}^{(2)}(\mathbf{x}_1^\alpha,\mathbf{x}_2^\eta) g_\eta(\mathbf{x}_2^\eta) d\mathbf{x}_2^\eta \\ &- \sum_{\eta=1,\eta\neq\alpha}^{s} \Delta_{\mathbf{x}_1^\alpha} \ln\left(\int_\Omega \tilde{g}_{\eta;\alpha}(\mathbf{x}_1^\eta) \omega_{\alpha\eta}(\mathbf{x}_1^\alpha,\mathbf{x}_1^\eta) d\mathbf{x}_1^\eta \right)\end{aligned}$$

for the mean solvent density

$$\rho_\alpha^S(\mathbf{x}_1^\alpha) = \bar{\rho}_S g_\alpha(\mathbf{x}_1^\alpha) = \bar{\rho}_S g_\alpha^0(\mathbf{x}_1^\alpha)\exp(-u_\alpha(\mathbf{x}_1^\alpha)) \tag{27}$$

of solvent site $\alpha=1,\ldots,s$. We coin these equations the BGY3dM equations. They require the site-site pair distribution functions $g_{\alpha\eta}^{(2)}$ as input, which can be computed by the SS-BGY3dM equations beforehand. The intramolecular distribution functions $\omega_{\alpha\eta}$ are given by relation (19).

III. COMPUTATIONAL DETAILS

To solve the SS-BGY3dM and BGY3dM Eqs. [(18) and (27)] all functions are approximated on a regular grid with mesh size $h=L/m$, where L is the length of the domain $\Omega=[0,L]^3$ in one direction and m is the number of grid points in one direction. Although the site-site pair distribution functions are radial symmetric, the SS-BGY3dM equations are solved with full 3d resolution and the computed pair distribution functions are used as input for the BGY3dM equations. The nonlinear integro-differential Eqs. [(18) and (27)] are solved by a standard Picard iteration with simple mixing, i.e., with an additional damping. In every step of the iteration a Poisson problem with Dirichlet boundary conditions $u_{\alpha\gamma}(\partial\Omega)=0$ and $u_\alpha(\partial\Omega)=0$, $\alpha,\gamma=1,\ldots,s$ has to be solved. Since we compute the right hand sides of Eqs. (18) and (27) by means of Fourier transformations, the solution of the Poisson problem is computed by a diagonal scaling in Fourier space. The convolution integrals are solved by the convolution theorem, i.e.,

$$f * g = \mathcal{F}_3^{-1}(\mathcal{F}_3(f)\mathcal{F}_3(g)), \tag{28}$$

where $(*)$ denotes the convolution and the 3d Fourier transformation is defined as

$$\hat{g}(\mathbf{k}) := \mathcal{F}_3(g)(\mathbf{k}) = \int_{\mathbb{R}^3} g(\mathbf{x})\exp(-2\pi\iota\mathbf{k}\cdot\mathbf{x}) d\mathbf{x}, \tag{29}$$

$$\mathcal{F}_3^{-1}(\hat{g})(\mathbf{x}) = \int_{\mathbb{R}^3} \hat{g}(\mathbf{k}) \exp(2\pi \iota \mathbf{k} \cdot \mathbf{x}) d\mathbf{k}. \tag{30}$$

We employ the parallel FFT algorithm of the FFTW.[32] For simplicity, all appearing differential operators are applied in Fourier space. The Fourier transform of the intramolecular distribution functions $\omega_{\alpha\gamma}$ is analytically known,

$$\mathcal{F}_3(\delta(r_{12} - r_0^{\alpha\gamma}))(\mathbf{k}) = \frac{2}{|\mathbf{k}|} \sin(2\pi |\mathbf{k}| r_0^{\alpha\gamma}) r_0^{\alpha\gamma}, \tag{31}$$

and has not to be computed. For numerical stability, the divisions by the normalization functions have to be regularized. Hence, we replace the exact divisions by

$$\frac{\int_\Omega \mathbf{F}_{\alpha\eta}(\mathbf{r} - \mathbf{r}') \bar{g}^{(2)}_{\alpha\eta;\gamma}(\mathbf{r} - \mathbf{r}') \omega_{\gamma\eta}(\mathbf{r}') d\mathbf{r}'}{\max(n^\eta_{\alpha\gamma}(\mathbf{r}), \epsilon_\omega)}, \quad \forall\, \mathbf{r} \in \Omega \tag{32}$$

and

$$\frac{g_\alpha(\mathbf{x})}{\max(n^\gamma_\alpha(\mathbf{x}), \epsilon_g)}, \quad \forall\, \mathbf{x} \in \Omega, \tag{33}$$

with the regularization parameters $\epsilon_\omega = 10^{-1}$ and $\epsilon_g = 10^{-2}$, respectively.

A. The Coulomb potential

The application of the discrete Fourier transform requires periodicity of the involved functions with respect to the computational domain Ω. Assuming that the functions are of short range, i.e., they decay faster to zero than $1/r^3$ with the distance r in 3d space, we can simply choose the domain large enough, so that all functions are sufficiently small at the boundaries. The effect of the periodicity can then be eliminated by zero padding. Hence, the convolution integrals can be computed by discrete Fourier transforms for all short-range potentials as, e.g., the Lennard–Jones potential. However, if the Coulomb potential $1/|\mathbf{r}|$ is involved, which is not of short range, we cannot directly apply the discrete Fourier transform, since the force does not vanish at the boundaries and is not periodic with respect to the domain. Nevertheless, the convolution integrals can be computed by means of discrete Fourier transforms as follows: We consider the (notationally simplified) convolution integral

$$\int_\Omega \mathbf{F}(\mathbf{r}' - \mathbf{r}) g_\alpha(\mathbf{r}' - \mathbf{r}) g_\gamma(\mathbf{r}') d\mathbf{r}', \tag{34}$$

with the total force $\mathbf{F} = \mathbf{F}^{\mathrm{LJ}} + \mathbf{F}^C$ consisting of Lennard–Jones and Coulomb parts. We split the Coulomb force into short-range and long-range parts, such that the long-range part is smooth and therefore has fast decaying Fourier components. This is achieved by adding and subtracting a smooth charge distribution around the point charge at $\mathbf{r} = 0$. As in the particle-mesh-Ewald method[33] the charge distribution is chosen to be a Gaussian

$$\varrho(\mathbf{r}) = \left(\frac{G}{\sqrt{\pi}}\right)^3 \exp(-G^2 |\mathbf{r}|^2), \tag{35}$$

with a parameter G that determines the width of the function. The Coulomb potential (and its force) between two particles is divided by means of this shielding function into the following parts:

$$\begin{aligned} v^C(\mathbf{r}) &= v^{Cs}(\mathbf{r}) + v^{Cl}(\mathbf{r}) = q_\gamma \Phi^s(\mathbf{r}) + q_\gamma \Phi^l(\mathbf{r}), \\ \mathbf{F}^C(\mathbf{r}) &= \mathbf{F}^{Cs}(\mathbf{r}) + \mathbf{F}^{Cl}(\mathbf{r}) = -\nabla v^{Cs}(\mathbf{r}) - \nabla v^{Cl}(\mathbf{r}), \end{aligned} \tag{36}$$

with Φ^s and Φ^l the solutions of the Poisson equations

$$\begin{aligned} -\Delta \Phi^s &= \frac{1}{\epsilon_0} q_\alpha (\delta_3 - \varrho) \quad \text{in } \mathbb{R}^3, \\ -\Delta \Phi^l &= \frac{1}{\epsilon_0} q_\alpha \varrho \quad \text{in } \mathbb{R}^3, \end{aligned} \tag{37}$$

where q_α and q_γ denote the charges of the two particles and ϵ_0 is the dielectric constant. For the special choice of the function ϱ as in Eq. (35), the solution can be given analytically. We have

$$\begin{aligned} v^{Cs}(\mathbf{r}) &= \frac{1}{4\pi\epsilon_0} q_\alpha q_\gamma \frac{\mathrm{erfc}(G|\mathbf{r}|)}{|\mathbf{r}|}, \\ v^{Cl}(\mathbf{r}) &= \frac{1}{4\pi\epsilon_0} q_\alpha q_\gamma \frac{\mathrm{erf}(G|\mathbf{r}|)}{|\mathbf{r}|}, \end{aligned} \tag{38}$$

with erf the error function and erfc = 1 − erf the complementary error function. The complementary error function decays rapidly, whereas the error function decays as slowly as $1/|\mathbf{r}|$ and is smooth even at $|\mathbf{r}| = 0$.

We now want to use these properties and transform the integral (34) to make it efficiently computable. The total force $\mathbf{F}$ consists of a Lennard–Jones part $\mathbf{F}^{\mathrm{LJ}}$ and a Coulomb part $\mathbf{F}^C$ which can be split as discussed above. Hence, we transform the convolution integral (34) according to

$$\begin{aligned} &\int_\Omega \mathbf{F}(\mathbf{r}' - \mathbf{r}) g_\alpha(\mathbf{r}' - \mathbf{r}) g_\gamma(\mathbf{r}') d\mathbf{r}' \\ &\quad = \int_\Omega (\mathbf{F}(\mathbf{r}' - \mathbf{r}) g_\alpha(\mathbf{r}' - \mathbf{r}) - \mathbf{F}^{Cl}(\mathbf{r}' - \mathbf{r})) g_\gamma(\mathbf{r}') d\mathbf{r}' \\ &\quad + \int_\Omega \mathbf{F}^{Cl}(\mathbf{r}' - \mathbf{r}) g_\gamma(\mathbf{r}') d\mathbf{r}'. \end{aligned} \tag{39}$$

The first term can be treated as before, since the part in outer brackets is of short range. The second integral can be written as

$$\int_\Omega \mathbf{F}^{Cl}(\mathbf{r}' - \mathbf{r}) g_\gamma(\mathbf{r}') d\mathbf{r}' = -\nabla_{\mathbf{r}} \int_\Omega v^{Cl}(\mathbf{r}' - \mathbf{r}) g_\gamma(\mathbf{r}') d\mathbf{r}', \tag{40}$$

where we used again the fact that the derivative of a convolution can be shifted to its arguments. Now, we can take advantage of the rapid decay of the Fourier components of v^{Cl}. The Fourier transformation can even be given analytically as

$$\mathcal{F}_3(v^{Cl})(\mathbf{k}) = \frac{q_\alpha q_\gamma}{\epsilon_0} \exp\left(-\frac{\pi^2}{G^2}|\mathbf{k}|^2\right). \tag{41}$$

Hence, this integral can be computed easily by multiplying the Fourier components of v^{Cl} and g_γ and a subsequent inverse Fourier transformation of the result. The operator $\nabla_\mathbf{r}$ in front of the integral is eliminated in Fourier space by the inverse operator of the left-hand side of the SS-BGY3dM or BGY3dM equations.

Now, the convolution integrals with the long-range Coulomb force can efficiently be computed. But still, the assumption of periodicity of the involved functions introduces an error at the boundaries. Since we know that the distribution functions have to be of short range,[34] it follows that we have $g_\alpha^S(\partial\Omega) \approx 0$ for a domain Ω, which is large enough. Hence, we employ the boundary conditions $g_\alpha^S(\partial\Omega)=0$ and thereby $u_\alpha(\partial\Omega)=v_\alpha^{Cl}(\partial\Omega)$. Here, we restrict our discussion to the BGY3dM equations since the exact same considerations also hold for the SS-BGY3dM equations. To enforce these boundary conditions and to remove the errors due to the periodic boundary conditions, we can compute this error as the solution of

$$\Delta u_\alpha^* = 0 \quad \text{in } \Omega, \tag{42}$$

with

$$u_\alpha^*(\partial\Omega) = u_\alpha(\partial\Omega) - v_\alpha^{Cl}(\partial\Omega),$$

where $u_\alpha(\partial\Omega)$ are the boundary values of the previously computed solution of the BGY3dM equations for particle type α. Subtraction of the solution u_α^* from u_α leads exactly to the solution of the BGY3dM equation with the desired Dirichlet boundary condition $\bar{u}_\alpha(\partial\Omega)=v_\alpha^{Cl}(\partial\Omega)$ where $\bar{u}_\alpha = u_\alpha - u_\alpha^*$.

The solution of Eq. (42) is computed by a simple finite-difference scheme with a seven-point stencil. The resulting systems of equations can be solved by any iterative method. For convenience we choose the GMRES method with block Jacobi preconditioning as it is implemented in PETSc.[35–37] The computational effort necessary to solve this problem is very small compared to the costs for solving the BGY3dM equations, since the repeated solution of Eq. (42) becomes very efficient at later steps of the Picard iteration. Then, subsequent iterates u_α^l, with l denoting the iteration number, only differ a little at the boundaries and we have $u_\alpha^{*l} \approx u_\alpha^{*l-1}$. Therefore, only few GMRES iterations are sufficient to compute the solution u_α^{*l} up to a prescribed accuracy.

IV. RESULTS

To validate the BGY3dM method for the computation of solvent densities around arbitrary solutes, we compare results computed by the BGY3dM method with results obtained from MD simulations. For this, we employ the HCl-like models of Hirata *et al.*[38] as solvent. They are two site models where the intermolecular interaction is described by a combination of Lennard–Jones and Coulomb forces. Hence, the total potential between two particles of species α and γ with α, γ=H, Cl can be written as

TABLE I. Parameter values for the HCl-like model solvents from Hirata *et al.* (Ref. 38).

$m_{Cl}=35.453u$	$m_H=1.008u$
$q_{Cl}=-0.2e$	$q_H=0.2e$
$\epsilon_{Cl}=0.5143$ kcal/mol	$\epsilon_H=0.0397$ kcal/mol
$\sigma_{Cl}=3.353$ Å	$\sigma_H^1=2.735$ Å, $\sigma_H^2=0.4$ Å

$$v_{\alpha\gamma}^I(r) = v_{\alpha\gamma}^{\mathrm{LJ}}(r) + v_{\alpha\gamma}^C(r) = 4\epsilon_{\alpha\gamma}\left(\left(\frac{\sigma_{\alpha\gamma}}{r}\right)^{12} - \left(\frac{\sigma_{\alpha\gamma}}{r}\right)^6\right) + \epsilon_C \frac{q_\alpha q_\gamma}{r}, \tag{43}$$

with $r=|\mathbf{x}^\alpha - \mathbf{x}^\gamma|$. The specific parameters for the HCl-like models can be found in Table I. The two models only differ in the value of σ_H, i.e., for the first model (HCl1) it is $\sigma_H^1 = 2.735$ Å and both particles have large Lennard–Jones spheres, whereas in the second model (HCl2) we have $\sigma_H^2 = 0.4$ Å and the Lennard–Jones sphere of the hydrogen is completely embedded inside that of the chloride atom.

The parameters for the Lennard–Jones potential are computed according to the Lorentz–Berthelot mixing rules

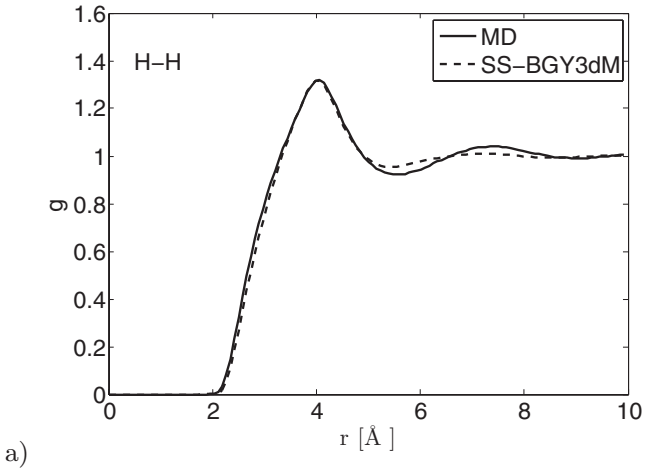
$$\epsilon_{\alpha\gamma} = \sqrt{\epsilon_\alpha \epsilon_\gamma}, \quad \sigma_{\alpha\gamma} = 0.5(\sigma_\alpha + \sigma_\gamma), \quad \alpha, \gamma = \mathrm{H}, \mathrm{Cl}. \tag{44}$$

The constant intramolecular distance between hydrogen and chloride atoms of the same molecule is set to $r_0^{\mathrm{HCl}} = 1.257$ Å. The molecular number density is chosen as $\bar{\rho}_S = 0.0018$ Å^{-3} with $\Omega=[-10\text{ Å}, 10\text{ Å}]^3$ and the temperature as $T=420$ K ($\beta=1.1989$). With this choice of unit system we additionally have $\epsilon_C \approx 331.84$ (kcal Å)/(mol e^2). We use $m=256$ and $m=128$ grid points in one direction for the solution of the SS-BGY3dM and the BGY3dM equations, respectively. The iteration of the (SS-)BGY3dM equations is stopped if the L_∞ norm of the difference of subsequent iterates is less than 10^{-2}. Note that this value for the stopping criterion is sufficient, because the model error of the (SS-)BGY3dM equations is larger anyway as we will see in the following section.

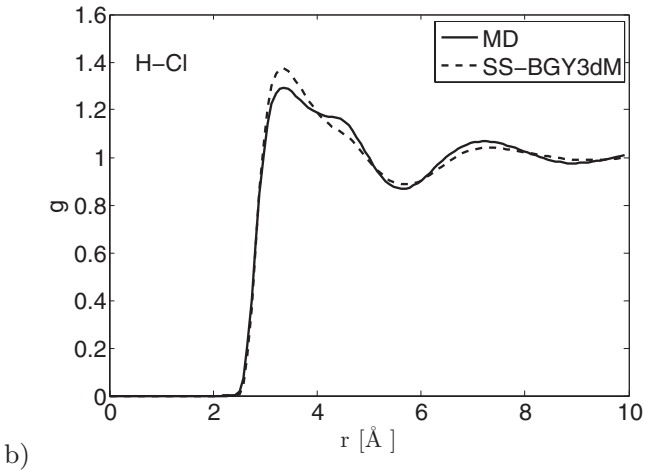
A. Site-site pair distribution functions

Now, we first compute the site-site pair distribution functions by means of the SS-BGY3dM Eqs. (18) and compare them to the pair distribution functions computed by a MD simulation of the pure solvent, see, e.g., Ref. 39 for details concerning the computation of pair distribution functions by MD. The results are shown in Figs. 1 and 2.

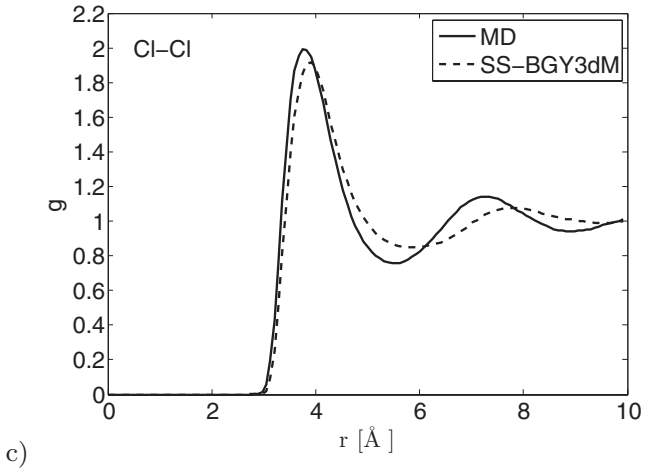
Note that the SS-BGY3dM equations are solved in three dimensions, but Figs. 1 and 2 only show the radial component of the computed site-site pair distribution functions. We compute three error values to quantitatively compare the results,



a)

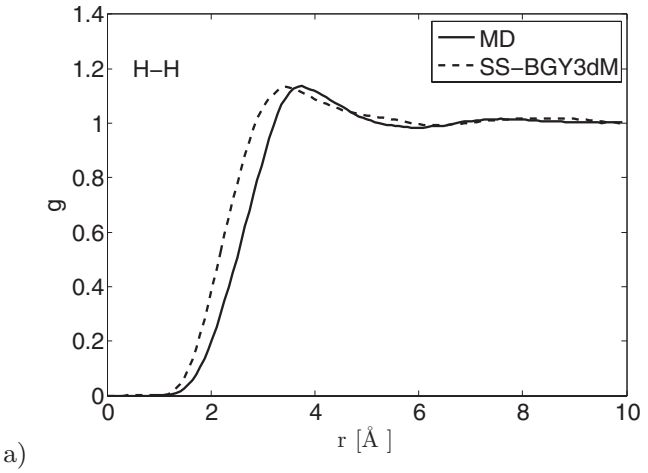


b)

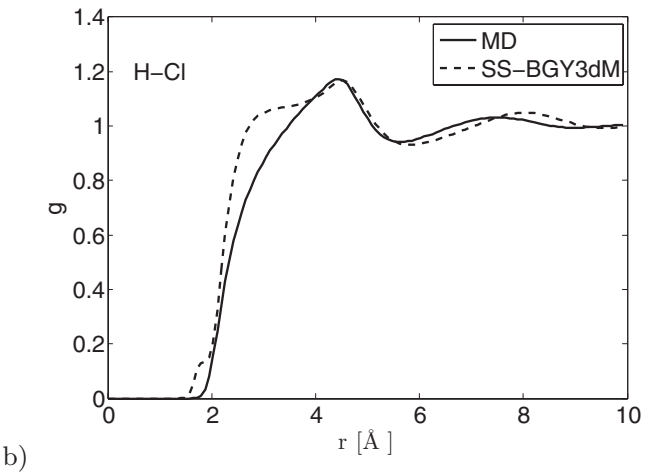


c)

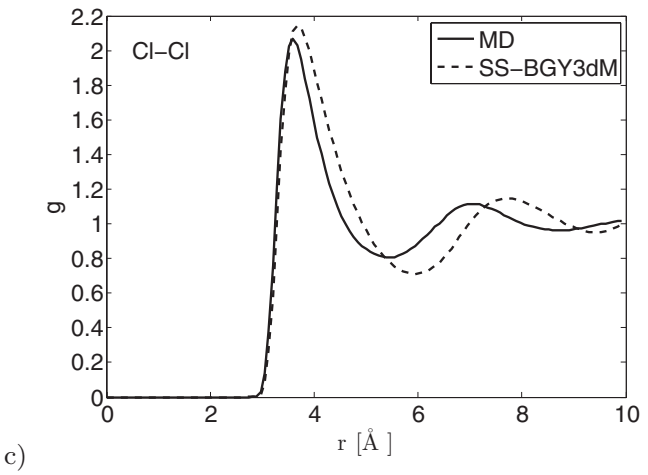
FIG. 1. Radial component of the site-site pair distribution functions for the HCl-like model (HCl1). Comparison between SS-BGY3dM and MD results.



a)



b)



c)

FIG. 2. Radial component of the site-site pair distribution functions for the HCl-like model (HCl2). Comparison between SS-BGY3dM and MD results.

$$e_{L_2^h} := \frac{1}{m^3}\left(\sum_{\mathbf{i}} |(g_h)_{\mathbf{i}} - (g_h^{\mathrm{MD}})_{\mathbf{i}}|^2\right)^{1/2},$$

$$e_{L_\infty^h} := \max_{\mathbf{i}} |(g_h)_{\mathbf{i}} - (g_h^{\mathrm{MD}})_{\mathbf{i}}|, \tag{45}$$

$$e_{\max} := |\max_{\mathbf{i}}(g_h)_{\mathbf{i}} - \max_{\mathbf{i}}(g_h^{\mathrm{MD}})_{\mathbf{i}}|,$$

where $(g_h)_{\mathbf{i}}$ denotes the respective solution of the SS-BGY3dM equations at grid point $\mathbf{i}$, $\mathbf{i} \in [0, m-1]^3$, and $(g_h^{\mathrm{MD}})_{\mathbf{i}}$ is the distribution function at grid point $\mathbf{i}$ computed with MD. To this end, the MD results of the pair distribution functions are interpolated on the 3d grid to compute the above errors. The values for the HCl-like model solvents can be found in Table II.

The apparent differences between the SS-BGY3dM and MD results in Figs. 1 and 2 are typical for any method employing the Kirkwood approximation. That is, the exact position and the height of the first peak do not exactly match those of the MD results except for the H–H distribution of HCl1. The frequency of the subsequent oscillation is too low. These errors are known to be a consequence of the two-

TABLE II. Comparison of SS-BGY3dM with MD results for the site-site pair distribution functions of the HCl-like models.

	MD	SS-BGY3dM			
	max g	max g	$e_{L_2^h}$	$e_{L_\infty^h}$	$e_{\max}$
H–H (HCl1)	1.32	1.32	3.549_{-6}	6.223_{-2}	0.00
H–Cl (HCl1)	1.29	1.37	4.385_{-6}	8.698_{-2}	0.08
Cl–Cl (HCl1)	1.99	1.93	1.491_{-5}	3.194_{-1}	0.06
H–H (HCl2)	1.14	1.13	6.893_{-6}	2.692_{-1}	0.00
H–Cl (HCl2)	1.17	1.18	7.516_{-6}	3.000_{-1}	0.01
Cl–Cl (HCl2)	2.07	2.48	2.896_{-5}	5.756_{-1}	0.41

particle superposition approximation, see e.g., Ref. 40. A comparison of the error values for the different site-site distribution functions of HCl1 reveals that their magnitude differs significantly. The L_2 and L_∞ errors of the Cl–Cl distribution function are about 3.5 times larger than the H–Cl errors. Hence, the quality of the approximation depends on the different potential parameters of the respective particle species. In this special example, the Cl atoms have a much stronger Lennard–Jones interaction, which obviously influences the quality of the solution in a negative sense. The comparison with the results for HCl2 uncovers another difficulty. All errors are increased for this model, which is due to the decreased value of σ_H considered for HCl2. This leads to a small Lennard–Jones sphere of the hydrogen atom which is completely embedded inside that of the chloride atom and results in a worse approximation of the SS-BGY3dM equations. Especially the H–H and H–Cl distribution functions of model 2 show major deficiencies in the prediction of the position of the first flank of the function.

Similar problems have been observed for the solution of the extended RISM equations of Hirata and Rossky.[38] For the H–Cl and Cl–Cl distribution functions of both HCl-like models the results of our SS-BGY3dM and the extended RISM model of Hirata and Rossky are very similar, although we have used a slightly increased temperature. The accuracy of the computed H–H distribution function is improved by the SS-BGY3dM model. In this case the SS-BGY3dM method predicts the correct height of the first peak for both the HCL1 and the HCL2 models and leads to an excellent overall accuracy for the HCL1 model.

We can conclude that the exact characteristics of all the site-site distribution functions are very difficult to approximate, as long as the approximations only comprise pair distribution functions. Obviously, the approximation of the SS-BGY3dM equations perform better for more similar particle species. However, note that the overall approximation of the presented method is still reasonable. All important features of the distribution functions of the HCl-like model solvent are reproduced. The general form and especially the modeling of the intramolecular bonds within the SS-BGY3dM equations is verified by the results and of the same quality as the results obtained by the extended RISM method.

B. Site density distributions

Next, we test the BGY3dM model with respect to the computation of solvent densities around a solute molecule. For this, we again employ the HCl-like model (HCl1) of Hirata *et al.*[38] already described above. First, a single HCl molecule is considered as the solute. It is placed symmetrically along the x_1 axis at the center of the simulation box. The site-site pair distribution functions of the pure solvent, which are required as input for the BGY3dM equations, are computed by the SS-BGY3dM model. All simulation parameters are chosen as described above. During the MD simulations the current site densities are interpolated on a 3d grid and averaged over all time steps to obtain the mean site density distributions. For this, a total of 3.2×10^8 MD time steps were necessary to reach a satisfactory level of convergence.

The computed site densities and their deviation are depicted in Fig. 3. The computed error quantities can be found in Table III. The MD results still show distinct fluctuations, but all features of the distribution functions have clearly developed. A comparison of the results of the BGY3dM model and MD shows a satisfying agreement. The low L_2 errors indicate a good overall approximation. The first peak and the subsequent oscillation pattern are reproduced with a sufficient accuracy considering the approximations involved in the model. The main difference can be observed at the location of the first peak of the distributions. All other errors are not resolved in the plots due to the fluctuations of the MD results. The L_∞ error is about 0.4 and 1.2 for the hydrogen and the chloride distributions, respectively, and is also located at the main peaks. The error of the chloride distribution function is considerably larger, as it was also the case for the Cl–Cl pair distribution functions, see the discussion above. Recall that the site-site pair distribution functions are required as input of the BGY3dM model. They are computed with the approximate SS-BGY3dM model. Hence, the approximation error enters twice: directly via the approximation involved in the BGY3dM model and by the use of the approximated site-site pair distribution functions computed with the SS-BGY3dM model.

Finally, we compare the site distribution functions of the HCl-like model solvent around a single hexane molecule ($CH_3(CH_2)_4CH_3$) as solute. The potential parameters for hexane are taken from the general-purpose force field OPLS.[41] The computed site densities and the difference between the BGY3dM and MD results are shown in Fig. 4. The computed error quantities can be found in Table III.

As for the HCl molecule as solute the overall agreement between the BGY3dM and MD results is good. The com-

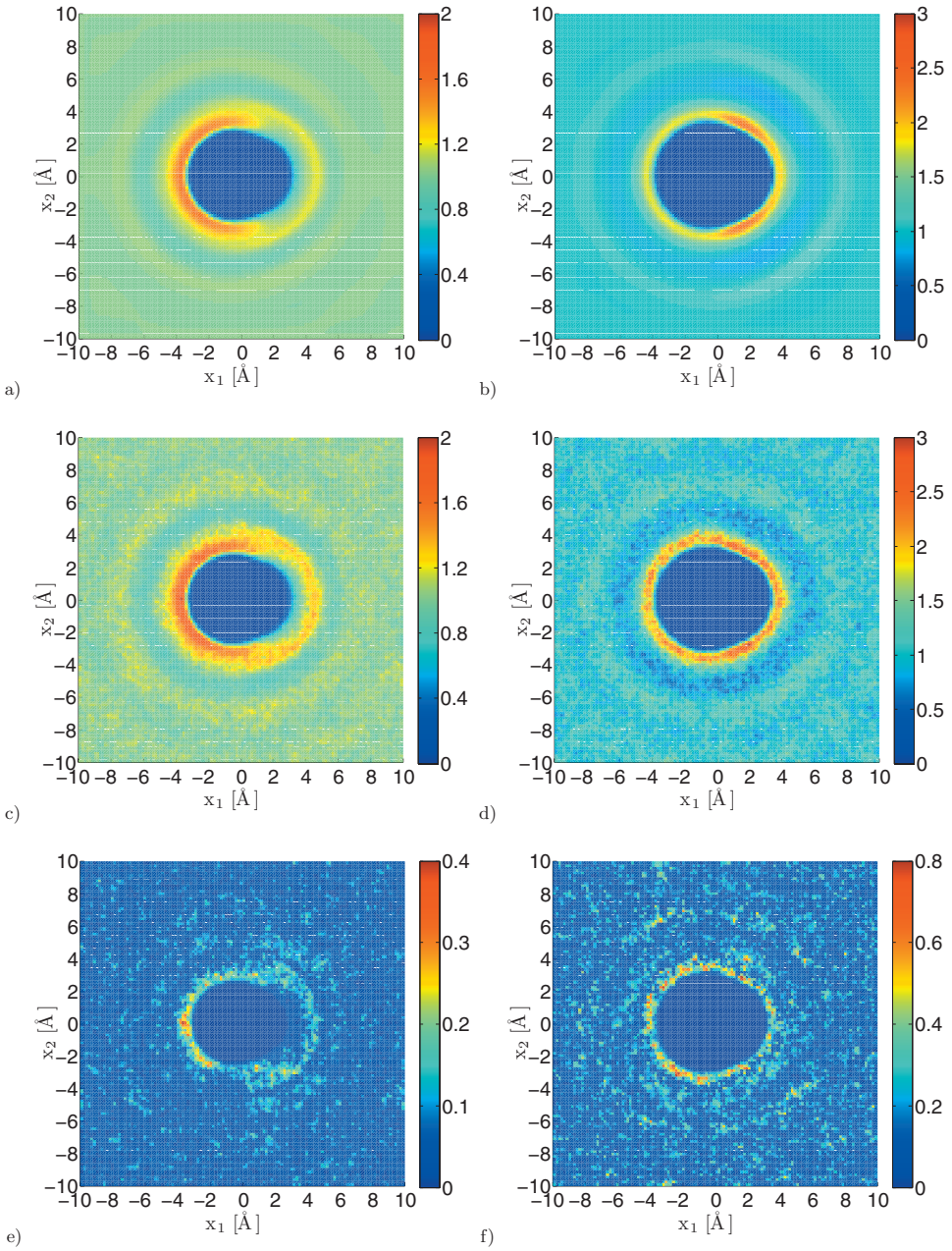


FIG. 3. (Color) Hydrogen [(a), (c), and (e)] and chloride [(b), (d), and (f)] distribution around a single HCl molecule at the $x_3=0$ plane computed with BGY3dM [(a) and (b)], MD [(c) and (d)], and the difference between BGY3dM and MD results [(e) and (f)].

puted error values are similar to those for HCl as solute. But again, the magnitude of the maxima in the first shell of the approximated hydrogen and chloride distributions around the solute are considerably small. Moreover, the predicted magnitude of the oscillation pattern which follows the main peak is also small compared to the MD results. Nevertheless, the positions of the maxima of the first shell and the form of the subsequent oscillation pattern is reproduced well. We can conclude that the accuracy of the computed site distribution functions is nearly independent of the size and form of the considered solute as long as the interactions between solute and solvent are similar in strength. The atoms of the considered HCl and hexane models do not carry large partial charges directly exposed to the solvent. Hence, the approxi-

TABLE III. Comparison of BGY3dM with MD results for the site density distributions of the HCl-like solvent around a single HCl and a hexane molecule as solute.

	MD	BGY3dM			
	max g	max g	$e_{L_2^h}$	$e_{L_\infty^h}$	$e_{\max}$
H (solute: HCl)	1.96	1.56	3.594_{-5}	4.134_{-1}	0.39
Cl (solute: HCl)	3.39	2.38	9.671_{-5}	1.215_{+0}	1.00
H (solute: hexane)	2.42	1.92	4.802_{-5}	5.525_{-1}	0.50
Cl (solute: hexane)	4.49	3.96	9.880_{-5}	1.299_{+0}	0.52

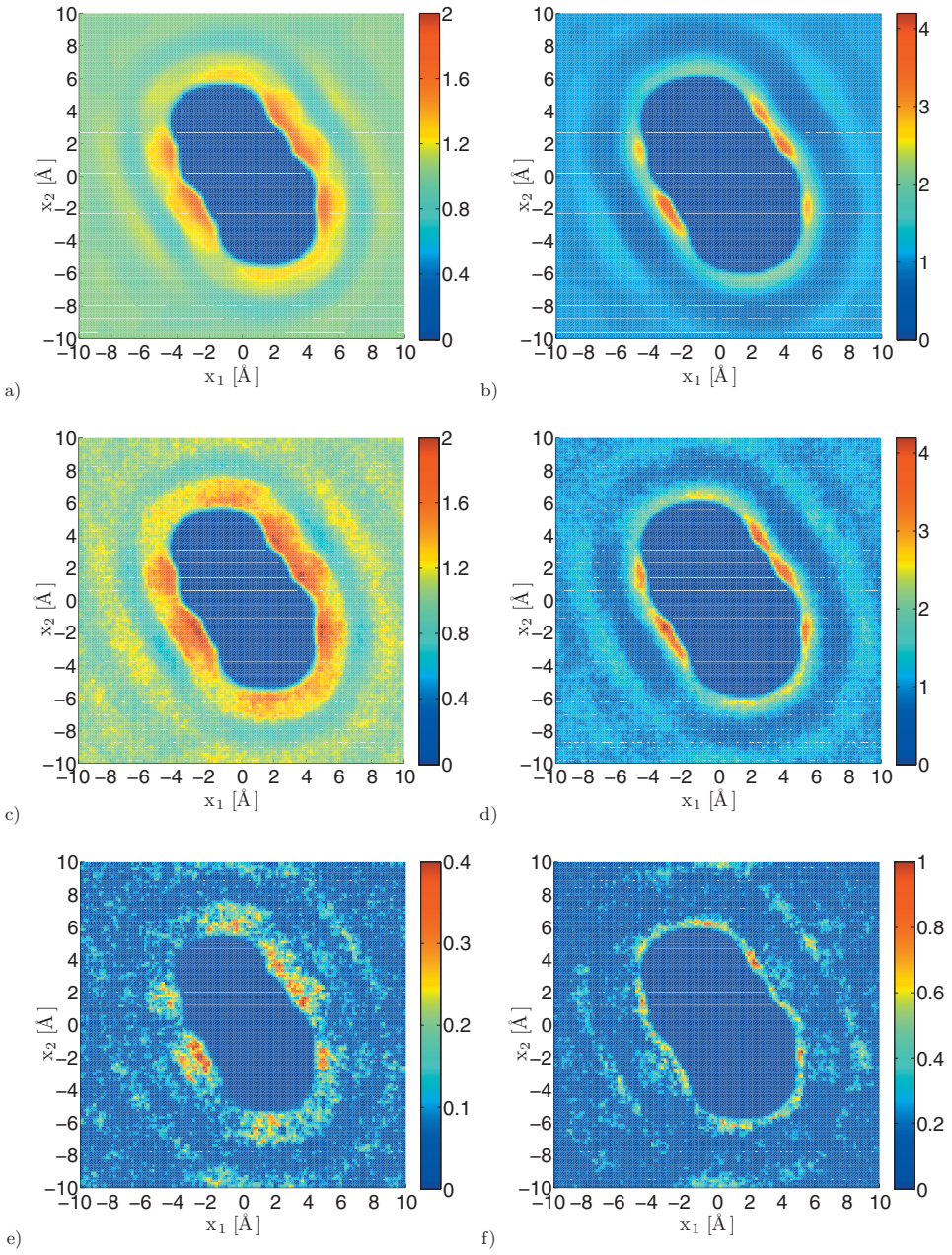


FIG. 4. (Color) Hydrogen [(a), (c), and (e)] and chloride [(b), (d), and (f)] distributions around a hexane molecule at the $x_3=0$ plane computed with BGY3dM [(a) and (b)], MD [(c) and (d)], and the difference between BGY3dM and MD results [(e) and (f)].

mation of the BGY3dM model yields similar results in both cases. This property is important to predict the error of a computed site density without actually comparing it to MD results.

C. Examples

We have seen that the BGY3dM model leads to a satisfying agreement between the computed site densities and the results of a MD simulation. Now, we present results obtained by the BGY3dM model for a more realistic fluid. For this, we consider carbon disulfide (CS_2) as solvent. Carbon disulfide is a colorless liquid which is mainly used to dissolve fats, rubber, resins and waxes, among other applications, see e.g., Ref. 42. The CS_2 molecule is linear and has no dipole moment. It is a nonpolar solvent. For our numerical computations we employ the model of Zhu *et al.*[43] As before, the functional form of the interaction potential is given as a sum of Lennard–Jones and Coulomb terms, see Eq. (43) with α, γ=C,S in this case.

Contrary to the HCl-like model solvent, the carbon di-

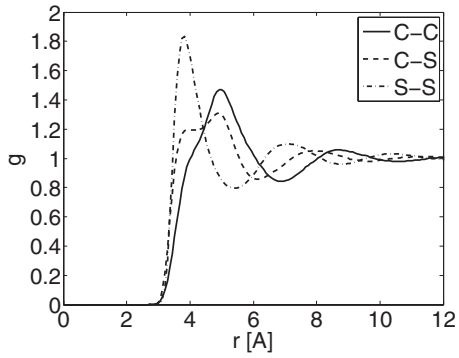


FIG. 5. The site-site pair distribution functions of carbon disulfide computed by MD.

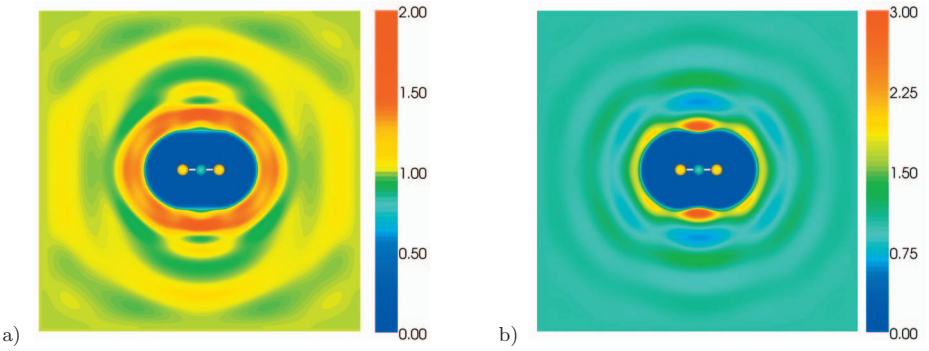


FIG. 6. (Color) Distribution functions of carbon disulfide around a CS_2 molecule at the $x_3=0$ plane. Carbon distribution (a) and sulfur distribution (b).

sulfide model is a three-site model. Since the two sulfur atoms of CS_2 are identical, we again have to compute three different site-site pair distribution functions and two different site distributions. The pair distribution functions are computed by a MD simulation of 80 CS_2 molecules at a number density of $\rho=0.01$ Å^{-3} and a temperature of $T=360$ K. The resulting C–C, C–S and S–S pair distributions are depicted in Fig. 5. They are used as input for the BGY3dM equations.

As a first example we compute the site distribution functions of carbon disulfide around a single CS_2 molecule as solute. The computational domain is set to $\Omega=[-14\text{ Å},14\text{ Å}]^3$. Figure 6 shows the site distributions at the $x_3=0$ plane. The carbon distribution exhibits a broad maximum around the entire solute molecule. This maximum results from a superposition of the van der Waals attraction modeled by the Lennard–Jones potential between the carbon atoms and the solute, and from the Coulomb attraction between the solvent carbon and the solute sulfur atoms. The solvent sulfur distribution shows a sharp peak around the solute carbon particle due to the strong Coulomb interaction between them. This can also be observed in Fig. 7, where the charge distribution is plotted. The charge distribution g_{charge} of carbon disulfide can be computed from the site distribution functions g_C and g_S by

$$g_{\text{charge}} = q_C g_C + 2 q_S g_S, \tag{46}$$

with q_α the charge of site $\alpha=C,S$. As can be expected, a closed band of high sulfur density evolves around the solute carbon whereas the solvent carbon is more likely to be found next to the solute sulfur atoms.

Next, we consider methanol as solute in carbon disulfide as solvent. Methanol is the simplest alcohol and has the chemical formula CH_3OH. It is a colorless, highly flammable liquid used as a petrol additive, solvent, or as antifreeze.[44] Due to the alcohol specific OH group, methanol is a polar molecule. The oxygen and hydrogen atoms carry strong opposed charges. We again employ the OPLS force field[41] for the parameter set of methanol. To this end, the hydrogen particle of the OH group is modeled as a charge carrying site without Lennard–Jones interaction. However, for the numerical stability of the BGY3dM equations a hard core potential is required at the position of any atom. Hence, we introduce Lennard–Jones parameters for the oxygen bonded hydrogen and choose $\sigma_H=3.4$ Å and $\epsilon_H=0.03$ kcal/mol. The high value of σ_H includes an empirical correction and ensures a stable convergence of the BGY3dM equations. We observed that the carbon density of the CS_2 solvent is overestimated in the neighborhood of strong positively charged particles as the hydrogen atom. This is partly due to a lack of an intramolecular coupling in the BGY3dM model where the two sulfur sites are incorporated without considering their relative position to the carbon site. This is a three-body effect that is neglected by the n-level Kirkwood approximation. Taking the three-body effect into account would lower the carbon density next to the hydrogen atom, because sulfur has a low density in the vicinity of positive charges. In order to compensate for this missing three-body effect, we choose the high value of $\sigma_H=3.4$ Å as an empirical correction.

Figure 8 shows the carbon and sulfur distributions around methanol at the $x_3=0$ plane. Here, the methanol molecule is depicted 2 Å above the plane for visualization purposes. It is obvious that the strong Coulomb interaction strongly influences the behavior of the distribution functions. The negatively charged solvent carbons are more likely to be found in the vicinity of the positive solute hydrogens, whereas the solvent sulfur atoms are dominantly attracted by the negative oxygen site. The plots of the charge distribu-

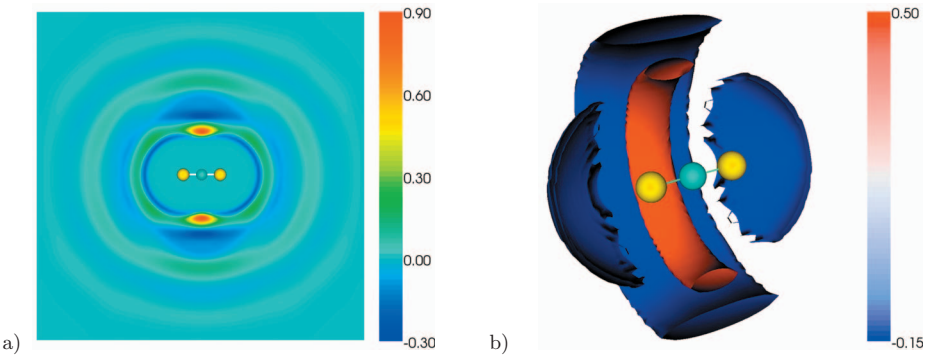


FIG. 7. (Color) Charge distribution of carbon disulfide around a CS_2 molecule. Cut at the $x_3=0$ plane (a) and isosurface plot (b).

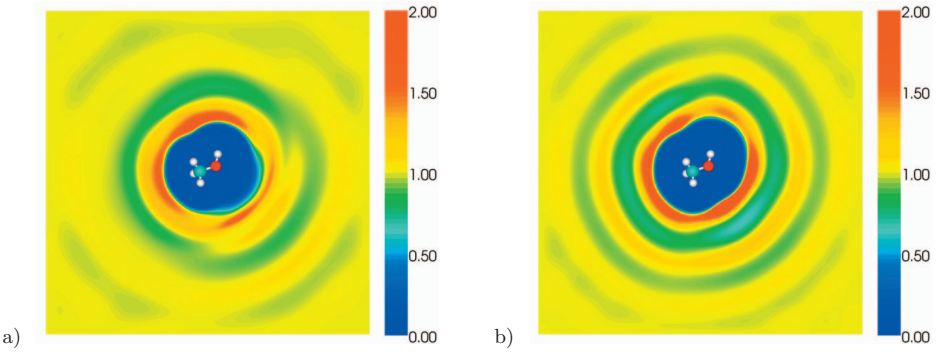


FIG. 8. (Color) Distribution functions of carbon disulfide around a methanol molecule at the $x_3=0$ plane. Carbon distribution (a) and sulfur distribution (b).

tions in Fig. 9 also uncover the negatively charged cloud behind the strong positive sulfur peak next to the solute oxygen atom. This charge minimum forms partly due to the intramolecular bond between carbon and sulfur, but also due to the intermolecular attraction of the different solvent sites. The whole picture reveals the well-known fact that charges tend to neutralize each other. Hence, the net forces on a particle in a fluid at equilibrium are exerted only by nearby particles although the long-range Coulomb potential is involved.

V. CONCLUSIONS

We have presented the BGY3dM model for the approximation of solvent densities around solutes of arbitrary shape. The model is directly derived from the YBG hierarchy and comprises the Kirkwood superposition approximation as closure relation for the intermolecular interactions. The intramolecular terms were derived to model rigid bonds by taking the limit of an infinite restoring force between two bonded particles. This way, the solvent molecules are represented as rigid bodies. Since the Kirkwood approximation is not appropriate for terms including intramolecular interactions, we employ a slightly simplified version of the NSSA of Taylor and Lipson[17] for these terms.

Beside the short-range Lennard–Jones potential, we also considered the long-range Coulomb interaction. For this, we introduced a splitting of the Coulomb potential into a singular short-range part and a smooth long-range part. The short-range part is processed in exactly the same way as the Lennard–Jones potential. The long-range part has fast decaying analytic Fourier components which are therefore directly dealt with in Fourier space. Nevertheless, the inverse Fourier transform of this long-range part leads to undesirable boundary conditions that have to be corrected. The correction comprises the solution of an additional Laplace problem which can efficiently be solved by a finite-difference scheme with an iterative GMRES solver. Finally, we also derived the SS-BGY3dM equations to compute the site-site pair distribution functions of the pure solvent which are required as input of the BGY3dM model.

Since both the YBG hierarchy and the Ornstein–Zernike equation are exact, the difference between our (SS-)BGY3dM and the RISM models originates in the different numerical methods to solve the equations and, more importantly, in the application of different approximations to close the equations. A comparison between both methods which employ the Kirkwood and the HNC closure, respectively, revealed that the accuracy of the results is nearly identical for the considered model solvent. However, since RISM based methods for the computation of solvent densities have been developed for about 25 years one cannot expect that the (SS-)BGY3dM method is of comparable quality for realistic solvents, yet. Nevertheless, it enables the use of different approximations and empirical corrections that could prove advantageous in the future.

A comparison of the results computed by the (SS-)BGY3dM model and by MD revealed a good overall performance of our method. All important characteristics of the site-site pair distribution functions and the site density distributions are reproduced. Hence, the general form of the (SS-)BGY3dM model including the modeling of the intramolecular bonds is consistent with the results. Nevertheless, the involved approximations have to be further modified to reach an improved level of accuracy at realistic temperatures and for stronger interactions. A promising approach would be the introduction of an empirical correction specific

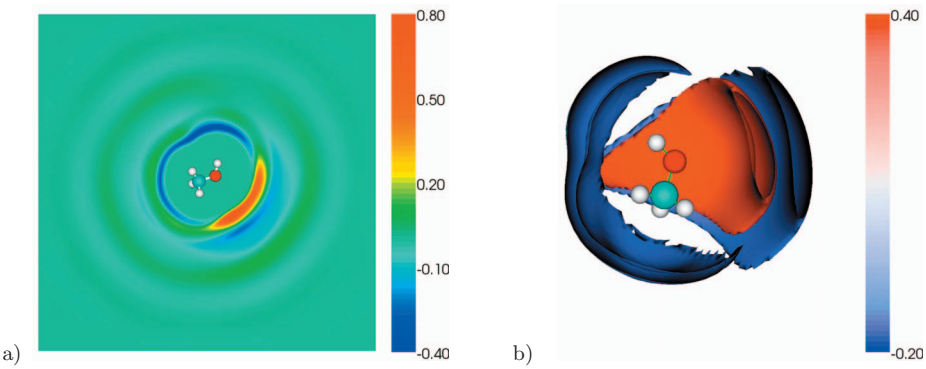


FIG. 9. (Color) Charge distribution of carbon disulfide around a methanol molecule. Cut at the $x_3=0$ plane (a) and isosurface plot (b).

for a certain solvent. This would be similar to the empirical bridge functions as they are employed by Du *et al.*[14] and Kovalenko and Hirata[10] for water as solvent. A better accuracy could also be gained by employing the optimal superposition approximation of the intramolecular terms as it has been derived by Attard.[18] This, however, requires an iterative solution of an additional system of equations for the functions to be superposed and would therefore increase the computational costs.

We conclude that our results are promising. The BGY3dM model is able to reproduce the important features of the site density distributions around an arbitrary solute. All solvent effects and thermodynamic properties of the solute-solvent system can then be computed from the site density distributions by thermodynamic integration as described by Du *et al.*[14] To consider more realistic solvents, such as water, empirical corrections of the approximations are currently under development.

ACKNOWLEDGMENTS

This work was supported by the "Deutsche Forschungsgemeinschaft" through the SFB 611 "Singuläre Phänomene und Skalierung in Mathematischen Modellen."

[1] *New Algorithms for Macromolecular Simulation*, Lecture Notes in Computational Science and Engineering Vol. 49, edited by B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, C. Schütte, and R. Skeel (Springer, Heidelberg, 2006).
[2] B. Roux and T. Simonson, Biophys. Chem. **78**, 1 (1999).
[3] B. Roux, in *Computational Biophysics*, edited by O. Becker, A. D. MacKerrel, B. Roux, and M. Watanabe (Dekker, New York, 2001).
[4] M. Ikeguchi and J. Doi, J. Chem. Phys. **103**, 5011 (1995).
[5] D. Beglov and B. Roux, J. Chem. Phys. **103**, 360 (1995).
[6] A. Kovalenko and F. Hirata, Chem. Phys. Lett. **290**, 237 (1998).
[7] A. Kovalenko and F. Hirata, J. Phys. Chem. B **103**, 7942 (1999).
[8] A. Kovalenko and F. Hirata, J. Chem. Phys. **112**, 10391 (2000).
[9] A. Kovalenko and F. Hirata, J. Chem. Phys. **112**, 10403 (2000).
[10] A. Kovalenko and F. Hirata, J. Chem. Phys. **113**, 2793 (2000).
[11] A. Kovalenko and F. Hirata, J. Chem. Phys. **113**, 9830 (2000).
[12] A. Kovalenko and T. N. Truong, J. Chem. Phys. **113**, 7458 (2000).
[13] D. Beglov and B. Roux, J. Phys. Chem. B **101**, 7821 (1997).
[14] Q. Du, D. Beglov, and B. Roux, J. Phys. Chem. **104**, 796 (2000).
[15] D. Chandler and H. C. Andersen, J. Chem. Phys. **57**, 1930 (1972).
[16] B. C. Eu and H. H. Gan, J. Chem. Phys. **99**, 4084 (1993).
[17] M. P. Taylor and J. E. G. Lipson, J. Chem. Phys. **100**, 518 (1994).
[18] P. Attard, J. Chem. Phys. **102**, 5411 (1995).
[19] H. H. Gan and B. C. Eu, J. Chem. Phys. **99**, 4103 (1993).
[20] H. H. Gan and B. C. Eu, J. Polym. Sci., Part B: Polym. Phys. **33**, 2319 (1995).
[21] H. H. Gan and B. C. Eu, AIChE J. **42**, 2960 (1996).
[22] H. H. Gan and B. C. Eu, J. Chem. Phys. **110**, 3235 (1999).
[23] M. P. Taylor and J. E. G. Lipson, J. Chem. Phys. **102**, 2118 (1995).
[24] M. P. Taylor and J. E. G. Lipson, J. Chem. Phys. **102**, 6272 (1995).
[25] M. P. Taylor and J. E. G. Lipson, J. Chem. Phys. **104**, 4835 (1996).
[26] M. P. Taylor, J. L. Mar, and J. E. G. Lipson, J. Chem. Phys. **106**, 5181 (1997).
[27] M. P. Taylor, J. Luettmer-Strathmann, and J. E. G. Lipson, J. Chem. Phys. **114**, 5654 (2001).
[28] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[29] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, London, 1986), 2nd ed..
[30] S. Sokolowski and J. Fischer, Mol. Phys. **70**, 1097 (1990).
[31] A. Singer, J. Chem. Phys. **121**, 3657 (2004).
[32] M. Frigo and S. G. Johnson, Proc. IEEE **93**, 216 (2005).
[33] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[34] A. Alastuey and P. A. Martin, J. Stat. Phys. **39**, 405 (1985).
[35] S. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang, PETSc Web page, 2001 (http://www.mcs.anl.gov/petsc).
[36] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith, in *Modern Software Tools in Scientific Computing*, edited by E. Arge, A. M. Bruaset, and H. P. Langtangen (Birkhäuser, Boston, 1997), pp. 163–202.
[37] S. Balay, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang, Technical Report No. ANL-95/11, Revision 2.1.5, Argonne National Laboratory, 2004.
[38] F. Hirata, B. M. Pettitt, and P. J. Rossky, J. Chem. Phys. **77**, 509 (1982).
[39] M. Griebel, S. Knapek, and G. Zumbusch, *Numerical Simulation in Molecular Dynamics. Numerics, Algorithms, Parallelization, Applications*, Texts in Computational Science and Engineering (Springer, New York, 2007).
[40] P. Attard, Mol. Phys. **74**, 547 (1991).
[41] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).
[42] Carbon Disulfide Fact Sheet 2007 (http://www.npi.gov.au/database/substance-info/profiles/18.html).
[43] S.-B. Zhu, J. Lee, and G. W. Robinson, Mol. Phys. **65**, 65 (1988).
[44] Methanol Fact Sheet, 2007 (http://www.npi.gov.au/database/substance-info/profiles/54.html).