# A molecular Density Functional Theory for associating fluids in 3D geometries

Antoine Barthes, Thomas Bernet, David Grégoire, Christelle Miqueu

# A molecular Density Functional Theory for associating fluids in 3D geometries

Antoine Barthes,[1] Thomas Bernet,[2] David Grégoire,[1,3] and Christelle Miqueu[1]

[1)]*Universite de Pau et des Pays de l'Adour, E2S UPPA, CNRS, LFCR, Anglet, France*

[2)]*Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, Institute for Molecular Science and Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

[3)]*Institut Universitaire de France, Paris*

(*Electronic mail: christelle.miqueu@univ-pau.fr)

(Dated: 11 December 2023)

A new free-energy functional is proposed for inhomogeneous associating fluids. The general formulation of Wertheim's thermodynamic perturbation theory is considered as the starting point of the derivation. We apply the hypotheses of the statistical associating fluid theory in the classical density functional theory (DFT) framework to obtain a tractable expression of the free-energy functional for inhomogeneous associating fluids. Specific weighted functions are introduced in our framework to describe association interactions for a fluid under confinement. These weighted functions have a mathematical structure similar to the weighted densities of the fundamental-measure theory (*i.e.*, they can be expressed as convolution products), such that they can be efficiently evaluated with Fourier transforms in a 3D space. The resulting free-energy functional can be employed to determine the microscopic structure of inhomogeneous associating fluids of arbitrary 3D geometry. The new model is first compared with Monte Carlo simulations and previous versions of DFT for a planar hard wall system in order to check its consistency in a 1D case. As an example of application in a 3D configuration, we investigate then the extreme confinement of an associating hard-sphere fluid inside an anisotropic open cavity with a shape that mimics a simplified model of zeolite. Both the density distribution and the corresponding molecular bonding profile are given, revealing complementary information to understand the structure of the associating fluid inside the cavity network. The impact of the degree of association on the preferential positions of the molecules inside the cavity is investigated as well as the competition between association and steric effect on adsorption.

*List of acronyms*

| | |
|---|---|
| TPT | thermodynamic perturbation theory |
| SAFT | statistical associating fluid theory |
| EoS | equation of state |
| PDF | pair distribution function |
| ref. | reference |
| HS / hs | hard sphere |
| RDF | radial distribution function |
| DFT | density functional theory |
| FMT | fundamental-measure theory |
| iSAFT | interfacial statistical associating fluid theory |
| aFMT | fundamental-measure theory for association |
| aWDA | weighted density approximation for association |
| assoc. | association |
| dist. | distance |
| ang. | angular |
| ext. | external |
| eq. | equilibrium |
| id. | ideal |
| ex. | exces |
| WBII | White Bear version mark II of FMT |
| RHNC | reference hypernetted-chain approach |
| CG | coarse grained |
| FFT | fast Fourier transform |
| LTA | Linde Type A |

## I. INTRODUCTION

Within the domain of thermodynamics, associating fluids constitute a unique class of substances known for their distinctive ability to form oriented interactions, *e.g.*, hydrogen bonds and strong polar interactions. When associating fluids are confined in porous media, a complex interplay ensues. The characterization of the fluid-solid interfaces and effects of confinement is

not only of interest for fundamental research, but is also of great importance in a variety of applied fields, including nanotechnology, heterogeneous catalysis, gas storage and separation, and environmental science[1–3].

In the 80's, Wertheim developed in its famous thermodynamic perturbation theory (TPT) a rigorous framework that models highly directional attractive intermolecular interactions capable of describing the thermodynamic properties of inhomogeneous associating fluids[4–7]. Within this pioneering approach, these attractive forces are mediated by acentrically located sites within each molecule, with the bonding constraints that two molecules can only form a single association with each other due to steric effects. A detailed discussion about Wertheim's framework can be found in the pedagogical review of Zmpitas *et al.*[8]. A few years later, Jackson *et al.*[9] restated Wertheim's TPT at first order, namely TPT1, in a framework for homogeneous fluids, such that the association contribution to the free energy could be introduced in the so-called SAFT equation-of-state (EoS) by Chapman *et al.*[10–12]. Multiple bonding sites are considered in a tractable way in the SAFT approach, by assuming the independence of the association interactions, such that the contributions to the free energy can be approximated as a sum over the different sites. Thus, estimating the association contribution to the Helmholtz free energy requires only to solve the so-called non-bonded fraction relation[9]. This latter is a key achievement of Wertheim's TPT and is formulated as a self-consistent equation involving a pair-correlation integral, that can be interpreted as the bonding probability, and requires the pair distribution function (PDF) of the reference inter-molecular potential $g^{ref.}(\boldsymbol{r}_1, \boldsymbol{r}_2)$. The PDF quantifies the extent by which the probability of finding a set of two molecules in the configuration $(\boldsymbol{r}_1, \boldsymbol{r}_2)$, *i.e.* the two-body density denoted as $\rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)$, deviates from the uncorrelated one-body densities $\rho(\boldsymbol{r}_1)$ and $\rho(\boldsymbol{r}_2)$, through the normalized relation $\left[ g(\boldsymbol{r}_1, \boldsymbol{r}_2) \equiv \rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \,/\, [\rho(\boldsymbol{r}_1)\,\rho(\boldsymbol{r}_2)] \right]$. In the TPT framework, the elaboration of non-ideal contributions from the PDF allows us to obtain tractable formulations both without explicit consideration of the partition function and also only expressed in terms of one-body density. Thus defined, the difficulty has been shifted in determining the PDF and one can use the individual properties of the target fluid to treat the pair correlation on a case-by-case basis. In the original work[9], as well as in our current work, a hard sphere (HS) fluid is considered as the reference non-associated fluid of the perturbation approach[4–7]. In the case of a uniform fluid, the correlation between two HS molecules only depends on the radial distance, designated as $g^{\mathrm{hs}}(r_{12})$, and is referred as the radial distribution function (RDF). An accurate description of the thermodynamic properties of a homogeneous HS fluid[13] along with the exact contact-value theorem[14] give access

to the particular value of the HS pair correlation at the contact distance $r_{12} = \sigma$, denoted here as $g^{\mathrm{hs}}_{\sigma}(\eta_\sigma)$, which only depends of the selected thermodynamic conditions, set by the packing factor $\eta_\sigma$. Because of the short-range intermolecular association, Jackson *et al.*[9] could approximate the RDF with the contact-value approach, and the resulting pair-correlation integral could be compacted by introducing the bonding volume (*i.e.*, the volume available to two sites for bonding). Thus defined, any configurations of association under the bonding constraints fixed in the TPT development can be considered with the SAFT formalism, but an analytical solution of the non-bonded fraction relation is only available for particular association schemes *e.g.*, schemes 1A, 2B, and 4C defined later. Subsequent studies on parameterization have enabled SAFT (the original formulation[11,12], as well as alternative parameterizations[15–21]) to be applied to the thermodynamic description of many real associative fluids (water, alcohols, acids, *etc.*) under bulk conditions[22–26].

In the more general case of inhomogeneous fluids, classical density functional theory[27] (DFT) has emerged as a powerful theoretical framework that provides the equilibrium density distributions of a fluid submitted to an external potential (such as the one induced by solid surfaces). In the DFT framework, the Helmholtz free energy, intrinsic to molecular interactions within the fluid, is a functional of the density distribution that can be approximated with a perturbative approach from a reference fluid, which is most often the HS fluid. A succession of theoretical achievements has resulted in an accurate thermodynamic description of inhomogeneous HS fluids; the most accurate being certainly the fundamental-measure theory[28–31] (FMT). Chapman[32] were the first to derive Wertheim's TPT into a tractable inhomogeneous form of the association free energy functional, the bulk limit of which was used in the development of SAFT briefly recalled above. However, this formulation is not directly applicable, as it requires both an inhomogeneous formulation of the PDF and a method for solving the integral equations arising from the inhomogeneity. In this respect, different paths have been followed in the literature, mainly using the bulk counterpart, not only as a guideline for constructing the inhomogeneous functional, but also as a homogeneous limit, such that the bulk parameterization can be employed in the inhomogeneous case. Segura *et al.*[33] proposed two methods to approximate the association contribution to the free-energy functional: a first one starting from the inhomogeneous form of Wertheim's theory, and a second one, starting from the bulk form (SAFT EoS) and extending it to an inhomogeneous scenario.

In the first method, Segura *et al.*[34] managed to derive Chapman's inhomogeneous expression into a formulation that can be applicable to adsorption on a planar solid surface by making two assumptions. The first one deals with the PDF and corresponds to the approximation that Jackson *et*

*al.*[9] employed in SAFT for homogeneous fluids. The second one relates to the treatment of the inhomogeneous functions (*e.g.* local density) involved in the bonding probability in the non-bonded fraction, whose final expression is given in the form of a 1D formulation. This is the starting point of the interfacial statistical associating fluid theory (iSAFT) developed by Chapman and coworkers[35–37]. Bymaster *et al.*[37] improved the inhomogeneous treatment of the PDF but did not propose any alternative resolution than Segura's approximation to solve the non-bonded fraction (as presented so far in the literature). In the second method, Segura *et al.*[33] used the bulk EoS as a starting point of the derivation, and replaced the bulk density by the weighted density introduced by Tarazona[38]. This method has also been adopted by Yu and Wu[39] who used Rosenfeld's FMT weighted densities (referred to as aFMT), and Camacho Vergara *et al.*[40] who used specific weighted density approximations (WDA) (referred to as aWDA). In contrast to iSAFT[35,36], the aFMT and aWDA formulations extend the bulk properties to an inhomogeneous framework at an earlier step of the derivation, thus making more difficult to adjudicate the treatment of the inhomogeneity.

A detailed inventory of these most common SAFT-based DFT versions for associating fluids can be found in the recent article of Camacho Vergara *et al.*[40] where they compared the iSAFT version of Bymaster *et al.*[37], aFMT, and aWDA approaches with molecular simulations for fluids in the neighborhood of a hard wall, for several thermodynamic conditions and several association schemes. Their study revealed that iSAFT shows slightly better agreement than aWDA with simulations for confined fluids when the association interactions are predominant. The aim of the present work is to propose a simpler and more general approach than the derivation of Chapman and coworkers[37,41] employed in iSAFT by: i) starting from the inhomogeneous framework of Wertheim[4–7] (as done in iSAFT); ii) following the recent formulations of Lurie-Cregg *et al.*[42] to improve the description of the inhomogeneous PDF for EoS-based DFT approaches; iii) applying the necessary approximations to find the homogeneous limit corresponding to the RDF used in SAFT[9]. The advantage of this new formulation is that it introduces weighted functions, specific to the range of association, that can be efficiently evaluated with Fourier transforms in a 3D space.

In Section II, the new approach for inhomogeneous associating fluids is presented. In particular, the intermolecular potential considered in our work is described in Section II A, the classical DFT framework is introduced in Section II B, and the new free-energy functional is detailed in Section II C. Results are shown in Section III for an associating fluid at contact with a hard planar wall (Section III A) and in an anisotropic 3D cavity (Section III B). In Section IV, we provide conclusions and closing remarks.

## II. THEORY

### A. Potential for inter-molecular repulsion and association interactions

We consider a single-component system of associating monomers. The total pairwise potential can be split into two contributions[9]

$$\phi^{\text{hs}}(\boldsymbol{r}_{12}) + \sum_{a,b} \phi_{ab}^{\text{assoc.}}(\boldsymbol{r}_{1a2b}) \tag{1}$$

where $\phi^{\text{hs}}$ is the HS potential, with $\boldsymbol{r}_{12}$ being the vector connecting the centers of the two particles 1 and 2. Denoting $\sigma$ the diameter of the particles,

$$\phi^{\text{hs}}(\boldsymbol{r}_{12}) = \begin{cases} \infty & \text{if} \quad r_{12} < \sigma \\ 0 & \text{if} \quad r_{12} \geq \sigma \end{cases} \tag{2}$$

where $r_{12} = |\boldsymbol{r}_{12}|$ is the center-center distance between two molecules. The second term of Eq. (1) is a sum over all association pair site potentials $\phi_{ab}^{\text{assoc.}}$ with $\boldsymbol{r}_{1a2b}$ being the vector connecting the position of site $a$ from molecule 1 to the position of site $b$ from molecule 2 (as shown in FIG. 1). Two standard formulations of $\phi_{ab}^{\text{assoc.}}$ exist. Association is simulated in the first simulation by means of a square-well potential centered on a site position and having a given cutoff $r_c$. Let us note $d$ the distance between on site position and the center of the molecule (FIG. 1(a)). Considering this distance formalism,

$$\phi_{\text{dist.},ab}^{\text{assoc.}}(\boldsymbol{r}_{1a2b}) = \begin{cases} -\varepsilon_{ab} & \text{if} \quad r_{ab} \leq r_c \\ 0 & \text{if} \quad r_{ab} > r_c \end{cases} \tag{3}$$

where $r_{ab} = |\boldsymbol{r}_{1a2b}|$ is the distance between association sites and $\varepsilon_{ab}$ is the association degree of interaction. A square-well potential which is centered on molecule positions is considered as a second formulation of association (FIG. 1(b)), for a given cutoff $r_{12,c}$, and the following angular parameters $(\theta_1, \theta_2)$ are introduced to define the site orientations. Let $\theta_1$ (respectively $\theta_2$) be the angle between $\boldsymbol{r}_{12}$ and the vector from the center of molecule 1 (respectively 2) to site $a$ (respectively $b$),

$$\phi_{\text{ang.},ab}^{\text{assoc.}}(\boldsymbol{r}_{1a2b}) = \begin{cases} -\varepsilon_{ab} & \text{if} \quad r_{12} \leq r_{12,c} \quad \text{and} \quad \theta_1 \leq \theta_{1,c} \quad \text{and} \quad \theta_2 \leq \theta_{2,c} \\ 0 & \text{if} \quad r_{12} > r_{12,c} \quad \text{or} \quad \theta_1 > \theta_{1,c} \quad \text{or} \quad \theta_2 > \theta_{2,c} \end{cases} \tag{4}$$

where $\theta_{1,c}$, $\theta_{2,c}$ are angular cutoffs. In order to compare our new DFT approach with existing

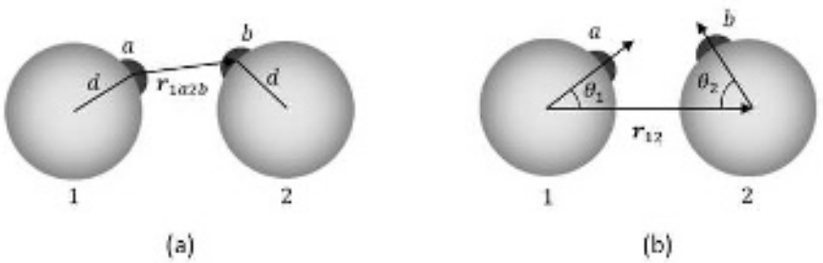



FIG. 1. Bonding description using distance parameters (a) and angular parameters (b).

Monte Carlo simulations[33,34,43], as well as previous DFT models[37,39,40], the angular formulation is employed in our current work, with $r_{12,c} = 1.05\sigma$ and $\theta_{1,c} = \theta_{2,c} = \theta_c = 27°$ for each site. Studies mentioned above considered three different schemes of association illustrated in FIG. 2 : scheme 1A for which there is one site A of each particle that can be bonded with another site ; in scheme 2B, two sites A and B are placed on the particle and we assume that only site A can be bonded with site B, *i.e.*, $\varepsilon_{AA} = \varepsilon_{BB} = 0$; scheme 4C consists of four associating sites A, B, C, and D where only AC, AD, BC, and BD are allowed to be bonded, *i.e.*, $\varepsilon_{AA} = \varepsilon_{AB} = \varepsilon_{BB} = \varepsilon_{CC} = \varepsilon_{CD} = \varepsilon_{DD} = 0$. The scheme 4C is typically used to model hydrogen bonding in the water molecule.

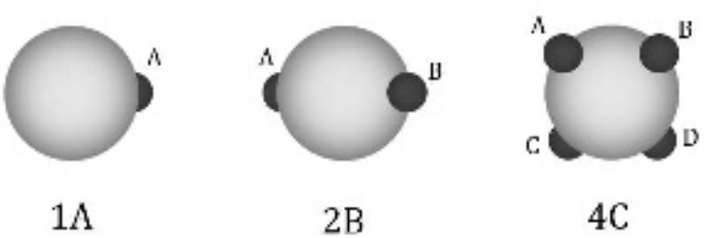



FIG. 2. Different schemes of association.

## B. DFT treatment for spherical associating monomer

Let us consider an ensemble of molecules in a volume $V$ at a given temperature $T$ and bulk chemical potential $\mu$ which is the grand canonical ensemble suitable for adsorption studies. Putting this system under an external field $V^{\text{ext.}}(\boldsymbol{r})$, the grand potential $\Omega[\rho]$ can be defined as a functional of the density profile $\rho(\boldsymbol{r})$ with the classical DFT framework[27]:

$$\Omega[\rho] = \mathscr{F}[\rho] + \int \rho(\boldsymbol{r}) \left( V^{\text{ext.}}(\boldsymbol{r}) - \mu \right) \mathrm{d}\boldsymbol{r} \tag{5}$$

where $\mathscr{F}[\rho]$ is the Helmholtz free-energy functional that does not depend explicitly on the external and bulk chemical potentials. The minimum of the grand potential is obtained for a unique density profile $\rho_{\text{eq.}}(\boldsymbol{r})$, such that

$$\left.\frac{\delta\Omega[\rho]}{\delta\rho(\boldsymbol{r})}\right|_{\rho(\boldsymbol{r})=\rho_{\text{eq.}}(\boldsymbol{r})} = 0 \tag{6}$$

and corresponds to the equilibrium state, for which the functional reduces to the grand-canonical potential $\Omega[\rho_{\text{eq.}}] = \Omega$. The free-energy functional $\mathscr{F}[\rho]$, intrinsic to a given pair potential model, can be split into an ideal part and an excess part, as $\mathscr{F} = \mathscr{F}^{\text{id.}} + \mathscr{F}^{\text{ex.}}$. The ideal part, corresponding to the kinetic energy, is known exactly from statistical thermodynamics[38]. The excess part, corresponding here to the interactions between associative HS, can be obtained from a thermodynamic expansion[4–7], $\mathscr{F}^{\text{ex.}} = \mathscr{F}^{\text{hs}} + \mathscr{F}^{\text{assoc.}}$, where the HS free-energy functional $\mathscr{F}^{\text{hs}}$ is defined as a reference, and the association free-energy functional $\mathscr{F}^{\text{assoc.}}$ as a perturbation. Once an expression of $\mathscr{F}^{\text{ex.}}$ is considered, one can minimize the grand potential to obtain the self-consistent density profile :

$$\rho(\boldsymbol{r}) = \rho\, \exp\left(\beta\mu^{\text{ex.}} - \frac{\delta\beta\mathscr{F}^{\text{ex.}}[\rho]}{\delta\rho(\boldsymbol{r})} - \beta V^{\text{ext.}}(\boldsymbol{r})\right) \tag{7}$$

where the density $\rho$ is the bulk limit of the one-body density $\rho(\boldsymbol{r})$ and $\mu^{\text{ex.}}$ the excess part of the bulk chemical potential. The functional derivative $\dfrac{\delta\mathscr{F}^{\text{ex.}}[\rho]}{\delta\rho(\boldsymbol{r})}$ estimates the response of the intrinsic Helmholtz free energy functional with respect to the density changes, *i.e.*, corresponds to the excess chemical potential for an inhomogeneous phase that can be decomposed as $\mu^{\text{ex.}}(\boldsymbol{r}) = \mu^{\text{hs}}(\boldsymbol{r}) + \mu^{\text{assoc.}}(\boldsymbol{r})$. This estimation is also known as the singlet-direct correlation function $c^{(1)}(\boldsymbol{r}) = -\mu^{\text{ex.}}(\boldsymbol{r})$. The bulk counterpart $\mu^{\text{ex.}}$ can be directly obtained by taking the homogeneous limit and thus the microscopic structure of a fluid submitted to an external potential can be completely determined only with an expression of $\mu^{\text{ex.}}(\boldsymbol{r})$.

Among the possible HS treatments, we consider here the HS free-energy functional of the FMT approach[28], with the White Bear version mark II (WBII)[30] (that corresponds to the Carnahan–Starling[13] EoS as a bulk limit), such that

$$\beta\mathscr{F}^{\text{hs}}[\rho] = \int \Phi^{\text{hs}}(n_\alpha(\boldsymbol{r}))\,\mathrm{d}\boldsymbol{r} \tag{8}$$

where $\beta = 1/(k_{\text{B}}T)$ ($k_{\text{B}}$ is the Boltzmann constant), and $\Phi^{\text{hs}}$ is the HS free-energy density, which is postulated to be a function of the weighted densities $n_\alpha(\boldsymbol{r})$ (where $\alpha$ denotes the nature of the weighting[30,44]). The generic form of a weighted density is a convolution product

$$n_\alpha(\boldsymbol{r}_1) = \int \rho\,(\boldsymbol{r}_2)\; p_\alpha^{\text{hs}}\,(\boldsymbol{r}_{12})\;\mathrm{d}\boldsymbol{r}_2 \tag{9}$$

where $p_\alpha^{\mathrm{hs}}$, that depends on the distance between the two positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, corresponds to the scalar- and vector-weight functions that describe the geometrical effects coming from the HS potential model[30,44]. Calculation of the inhomogeneous HS excess chemical potential functional can also be evaluated as a sum of convolution products, such that

$$\beta\mu^{\mathrm{hs}}(\boldsymbol{r}_1) = \frac{\delta\beta\mathscr{F}^{\mathrm{hs}}[\rho]}{\delta\rho(\boldsymbol{r}_1)} = \sum_\alpha \int \left.\frac{\partial\Phi^{\mathrm{hs}}}{\partial n_\alpha}\right|_{(\boldsymbol{r}_2)} p_\alpha^{\mathrm{hs}}(\boldsymbol{r}_{12})\,\mathrm{d}\boldsymbol{r}_2 \tag{10}$$

## C. Association contribution to the free-energy functional

### 1. *Wertheim's free-energy and chemical potential functionals*

The TPT1 Wertheim's association free energy[4–7] for a pure fluid of associating HS is given by

$$\beta\mathscr{F}^{\mathrm{assoc.}}[\rho] = \int \rho(\boldsymbol{r}_1) \sum_a \left[\ln X_a(\boldsymbol{r}_1) - \frac{X_a(\boldsymbol{r}_1)}{2} + \frac{1}{2}\right] \mathrm{d}\boldsymbol{r}_1 \tag{11}$$

as derived by Chapman[32] and Segura *et al.*[45]. $X_a(\boldsymbol{r}_1)$ is the fraction of a non-bonded site $a$ of a molecule at position $\boldsymbol{r}_1$,

$$X_a(\boldsymbol{r}_1) = \frac{1}{1 + \int \rho(\boldsymbol{r}_2)\, g^{\mathrm{hs}}(\boldsymbol{r}_1,\boldsymbol{r}_2) \sum_b X_b(\boldsymbol{r}_2)\, \bar{F}_{ab}(r_{12})\,\mathrm{d}\boldsymbol{r}_2} \tag{12}$$

where $g^{\mathrm{hs}}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is the HS PDF at positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, and $\bar{F}_{ab}(r_{12})$ is an average of the Mayer function $f_{ab}(\boldsymbol{r}_{1a2b}) = \Big[\exp(-\beta\phi_{ab}(\boldsymbol{r}_{1a2b})) - 1\Big]$ over the orientations $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ of molecules 1 and 2, respectively, such that

$$\bar{F}_{ab}(r_{12}) = \Big\langle f_{ab}(\boldsymbol{r}_{1a2b}) \Big\rangle_{\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2} \tag{13}$$

Because of the unweighted average over all orientations, the Mayer function only depends on $r_{12}$, and the considered one-body density $\rho(\boldsymbol{r}_1)$ and fraction $X_a(\boldsymbol{r}_1)$ do not depend on the orientations $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$. The function $\bar{F}_{ab}(r_{12})$ can be expressed with the distance association potential model[46],

$$\begin{aligned}\bar{F}_{\mathrm{dist.},ab}(r_{12}) = {} & F_{ab}\frac{(r_c + 2d - r_{12})^2\,(2r_c - 2d + r_{12})}{24\,d^2\,r_{12}} \\ & \times\Big(\Theta(2d + r_c - r_{12}) - \Theta(\sigma - r_{12})\Big)\end{aligned} \tag{14a}$$

or with the angular potential model[9],

$$\bar{F}_{\mathrm{ang.},ab}(r_{12}) = F_{ab}\frac{(1-\cos\theta_c)^2}{4}\Big(\Theta(r_{12,c} - r_{12}) - \Theta(\sigma - r_{12})\Big) \tag{14b}$$

where $F_{ab} = \exp(\beta\varepsilon_{ab}) - 1$ represents the association strength, and the Heaviside step function $\Theta$ (such that $\Theta(x) = 1$ if $x > 0$, and $\Theta(x) = 0$ if $x < 0$) is used to take into account the association boundaries.

Solving Eq. (7) requires a formulation of the chemical potential functional for the association contribution, $\mu^{\text{assoc.}}(\boldsymbol{r}_1) = \dfrac{\delta \mathscr{F}^{\text{assoc.}}}{\delta\rho(\boldsymbol{r}_1)}$. However, the derivative $\dfrac{\delta X_a(\boldsymbol{r}_1)}{\delta\rho(\boldsymbol{r}_1)}$ should not be expressed explicitly to avoid a self-consistent equation problem. This issue has been solved by Michelsen and Hendriks[47] for a homogeneous system, and by Chapman and coworkers[36,37] for an inhomogeneous system, such that

$$\begin{aligned}\beta\mu^{\text{assoc.}}(\boldsymbol{r}_1) = \sum_a \ln X_a(\boldsymbol{r}_1) - \frac{1}{2}\iint \rho(\boldsymbol{r}_2)\,\rho(\boldsymbol{r}_3)\sum_{a,b} X_a(\boldsymbol{r}_2) \\ \times X_b(\boldsymbol{r}_3)\,\frac{\delta g^{\text{hs}}(\boldsymbol{r}_2,\boldsymbol{r}_3)}{\delta\rho(\boldsymbol{r}_1)}\,\bar{F}_{ab}(r_{23})\,\mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{r}_3\end{aligned} \tag{15}$$

### 2. *Approximations on the pair distribution function (PDF)*

Exact expressions of the inhomogeneous PDF are unknown but they can be estimated by approximation methods. In a bulk phase, the PDF only depends on the radial distance $r_{12}$ and corresponds to the RDF. Jackson *et al.*[9] observed that for short distances beyond the contact value $\sigma$ between two molecules 1 and 2, the RDF of a homogeneous HS fluid decreases as $1/r_{12}^2$, so that it can be approximated as

$$g^{\text{hs}}(r_{12}\,;\,\eta_\sigma) \approx \frac{\sigma^2}{r_{12}^2} g_\sigma^{\text{hs}}(\eta_\sigma) \tag{16}$$

where $g_\sigma^{\text{hs}}(\eta_\sigma)$ is the contact value of the HS RDF, and $\eta_\sigma = \dfrac{\pi}{6}\rho\,\sigma^3$ is the bulk packing fraction. Dufal *et al.*[20,21] showed that this approximation fits well the HS RDF obtained with the reference hypernetted-chain (RHNC) integral equation theory[48], under the range of bonding. The function $g_\sigma^{\text{hs}}(\eta_\sigma)$ can be obtained by using the exact contact-value theorem for the Carnahan–Starling EoS[13], such that

$$g_\sigma^{\text{hs}}(\eta_\sigma) = \frac{1 - \eta_\sigma/2}{(1-\eta_\sigma)^3} \tag{17}$$

for a homogeneous HS fluid.

The treatment of the PDF for an inhomogeneous HS fluid is more complex than for its homogeneous equivalent because $g^{\text{hs}}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ depends both on the positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ (and not only on

the distance $r_{12}$), and on the density profile (and not only on a constant density). There is not currently a suitable PDF of an inhomogeneous HS fluid to be used in the association free-energy functional we develop here. Indeed, $\mathscr{F}^{\text{assoc.}}[\rho]$ has first to exactly correspond to the SAFT free energy $F^{\text{assoc.}}$ in the homogeneous limit, which requires that the PDF corresponds to the SAFT RDF (Eq. (16)) in the bulk limit. It excludes approaches based on external potential as for instance the Percus' method[49,50] which is unsuitable for SAFT-based free-energy functionals. Another strategy consists in extending the RDF to the inhomogeneous case. The simplest analytical approximation consists in replacing the PDF of the inhomogeneous system by the RDF of the homogeneous fluid at a specific density. Segura *et al.*[45] have already used this approximation, choosing the bulk density under the same thermodynamic conditions as the reference density, thus recovering the SAFT formulation of the non-bonded fraction in the bulk limit. However, this approximation is a rough estimate of the PDF since local densities near an interface may differ from the bulk value by several orders of magnitude. In addition, the free-energy functional should not depend explicitly on the bulk density to satisfy the Gibbs adsorption theorem[31]. Lurie-Gregg *et al.*[42] discussed the advantages and shortcomings of several analytical approximations for inhomogeneous PDF that transform RDF into an inhomogeneous framework using position-dependent reference densities. They introduced the so-called contact value approach to compute the PDF of an inhomogeneous HS fluid that is suitable for DFTs based on TPT and that only requires a low computational cost. We are inspired by their approach but apply the necessary approximations to find the homogeneous limit corresponding to the RDF used in the SAFT association term (Eq. (16)). We therefore use the mean-function approximation[42] so that

$$g^{\text{hs}}(\boldsymbol{r}_1,\boldsymbol{r}_2) \approx \frac{g^{\text{hs}}(r_{12}\,;\,\eta_\sigma(\boldsymbol{r}_1)) + g^{\text{hs}}(r_{12}\,;\,\eta_\sigma(\boldsymbol{r}_2))}{2} \tag{18}$$

in which the position-dependant RDF $g(r_{12}\,;\,\eta_\sigma(\boldsymbol{r}_1))$ and $g(r_{12}\,;\,\eta_\sigma(\boldsymbol{r}_2))$ are approximated as in (Eq. (16)). Hence,

$$g^{\text{hs}}(\boldsymbol{r}_1,\boldsymbol{r}_2) \approx \frac{\sigma^2}{2\,r_{12}^2}\left(g_\sigma^{\text{hs}}(\boldsymbol{r}_1) + g_\sigma^{\text{hs}}(\boldsymbol{r}_2)\right) \tag{19}$$

where we only consider the inhomogeneous HS PDF at the contact distance[51,52],

$$g_\sigma^{\text{hs}}(\boldsymbol{r}_1) = \frac{1-\eta_\sigma(\boldsymbol{r}_1)/2}{\left(1-\eta_\sigma(\boldsymbol{r}_1)\right)^3} \tag{20}$$

by analogy with Eq. (17), and where the bulk packing fraction $\eta_\sigma$ is replaced by the weighted density

$$\eta_\sigma(\boldsymbol{r}_1) = \int \rho(\boldsymbol{r}_2)\,\frac{\Theta(R\,\psi - r_{12})}{\psi^3}\,\mathrm{d}\boldsymbol{r}_2 \tag{21}$$

where $R = \sigma/2$ is the radius of the molecule, $\psi$ is a parameter defining the size of the weighting and the denominator factor $\psi^3$ is here to ensure that the homogeneous limit of the functional corresponds to the bulk packing fraction. One could optimize the parameter $\psi$ as it has already been done in some DFT work[53], but by simplicity we only consider here the traditional coarse-grained approach for which $\psi = 2$, and we define the coarse-grained weight $p^{\mathrm{cg}}(r_{12}) = \dfrac{\Theta(\sigma - r_{12})}{8}$ to compact the notation in the remaining derivation.

### 3. *Expression of the chemical potential functional with weighted functions*

With the separation of the variables $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in the approximated PDF (Eq. (19)), the pair correlation integrals in the non-bonded fraction (Eq. (12)) and the chemical potential functional (Eq. (15)) can be decomposed as a sum of one-center convolution products. The non-bonded fraction can then be expressed as

$$X_a(\boldsymbol{r}_1) = \frac{1}{1 + \sum_b \frac{1}{2}\left(g^{\mathrm{hs}}_\sigma(\boldsymbol{r}_1)\,\chi_{ba}(\boldsymbol{r}_1) + G^{\mathrm{assoc.}}_{ba}(\boldsymbol{r}_1)\right)} \tag{22}$$

where we introduce the weighted functions

$$\chi_{ba}(\boldsymbol{r}_1) = \int \rho(\boldsymbol{r}_2)\,X_b(\boldsymbol{r}_2)\,p^{\mathrm{assoc.}}_{ab}(\boldsymbol{r}_{12})\,\mathrm{d}\boldsymbol{r}_2 \tag{23}$$

and

$$G^{\mathrm{assoc.}}_{ba}(\boldsymbol{r}_1) = \int \rho(\boldsymbol{r}_2)\,X_b(\boldsymbol{r}_2)\,g^{\mathrm{hs}}_\sigma(\boldsymbol{r}_2)\,p^{\mathrm{assoc.}}_{ab}(\boldsymbol{r}_{12})\,\mathrm{d}\boldsymbol{r}_2 \tag{24}$$

which depend on the association weight, defined as follows

$$p^{\mathrm{assoc.}}_{ab}(\boldsymbol{r}_{12}) = \frac{\sigma^2}{|\boldsymbol{r}_{12}|^2}\,\bar{F}_{ab}(|\boldsymbol{r}_{12}|) \tag{25}$$

regardless the choice made to describe association bonding (angular or distance potential formulation). The chemical potential functional can be expressed as

$$\begin{aligned}\beta\mu^{\mathrm{assoc.}}(\boldsymbol{r}_1) = &\sum_a \ln X_a(\boldsymbol{r}_1) - \frac{1}{2}\sum_{a,b}\iint \rho(\boldsymbol{r}_2)\,X_a(\boldsymbol{r}_2)\,\rho(\boldsymbol{r}_3)\,X_b(\boldsymbol{r}_3)\\ &\times \left.\frac{\partial g^{\mathrm{hs}}_\sigma}{\partial\eta}\right|_{(\boldsymbol{r}_2)} p^{\mathrm{cg}}(\boldsymbol{r}_{12})\,p^{\mathrm{assoc.}}_{ab}(\boldsymbol{r}_{23})\,\mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{r}_3\end{aligned} \tag{26}$$

The last term in the right-hand side in Eq. (26) is simplified by introducing $\chi_{ba}$, and by defining the weighted function

$$\bar{G}^{\text{assoc.}}_{ab}(\boldsymbol{r}_1) = \int \rho(\boldsymbol{r}_2)\, X_a(\boldsymbol{r}_2)\, \left.\frac{\partial g^{\text{hs}}_{\sigma}}{\partial \eta}\right|_{(\boldsymbol{r}_2)} \chi_{ba}(\boldsymbol{r}_2)\, p^{\text{cg}}(r_{12})\, \mathrm{d}\boldsymbol{r}_2 \tag{27}$$

such that the chemical potential can be written as

$$\beta\mu^{\text{assoc.}}(\boldsymbol{r}_1) = \sum_a \ln X_a(\boldsymbol{r}_1) - \frac{1}{2}\sum_{a,b} \bar{G}^{\text{assoc.}}_{ab}(\boldsymbol{r}_1) \tag{28}$$

### 4. *Calculation of weighted functions*

Two types of weighted functions have been introduced in the new expression of the association functional proposed in our current work. Both have the same mathematical structure, similar to the weighted densities for HS (Eq. (9)), that can be interpreted as the convolution product of a local function $f$ and a given weight.

The first type of weighted function depends on the association weight $p^{\text{assoc.}}_{ab}$, for which we define the generic expression

$$n^{\text{assoc.}}(\boldsymbol{r}_1) = \int f(\boldsymbol{r}_2)\, p^{\text{assoc.}}_{ab}(\boldsymbol{r}_{12})\, \mathrm{d}\boldsymbol{r}_2 \tag{29}$$

by analogy with Eqs. (23–24). One can explicit $p^{\text{assoc.}}_{ab}$ by replacing $\bar{F}_{ab}(r_{12})$ in Eq. (25) with the expressions given in Eqs. (14a–14b), such that, for the distance formulation

$$\begin{aligned} p^{\text{assoc.}}_{\text{dist.},ab}(\boldsymbol{r}) = {} & \frac{\sigma^2}{r^2} F_{ab} \frac{(r_c + 2d - r)^2\,(2r_c - 2d + r)}{24\, d^2\, r} \\ & \times \Big(\Theta(2d + r_c - r) - \Theta(\sigma - r)\Big) \end{aligned} \tag{30a}$$

and for the angular formulation

$$\begin{aligned} p^{\text{assoc.}}_{\text{ang.},ab}(\boldsymbol{r}) = {} & \frac{\sigma^2}{r^2} F_{ab} \frac{(1 - \cos\theta_c)^2}{4} \\ & \times \Big(\Theta(r_{12,c} - r) - \Theta(\sigma - r)\Big) \end{aligned} \tag{30b}$$

where $r = |\boldsymbol{r}|$. The weighted function can be solved by using fast Fourier transforms (FFT) in a tridimensional space[54], as

$$n^{\text{assoc.}}(\boldsymbol{r}_1) = \text{FFT}^{-1}\left[\text{FFT}\left[f\right] \times \hat{p}^{\text{assoc.}}_{\text{ab}}(\boldsymbol{k})\right] \tag{31}$$

Although the main interest of Fourier transforms to solve convolution products in DFT approach comes from the use of the FFT algorithm[55], there are others favourable features. In the Fourier space, the weight $\hat{p}_{\mathrm{ab}}^{\mathrm{assoc.}}(\boldsymbol{k})$ can be computed independently of the density profile and thus has to be evaluated only once for the iterative method used to solve Eq. (7), for a given size of the system. Also, a Fourier transform implies by definition a periodicity of the weighted functions that is an asset if one desires to introduce periodic boundary conditions (the simplified model of zeolite that we present in Section III B showcases this feature).

The analytical expressions of the Fourier transform of the association weight $\hat{p}_{\mathrm{ab}}^{\mathrm{assoc.}}(\boldsymbol{k})$ for the distance and angular formulations are

$$\hat{p}^{\mathrm{assoc.}}_{\mathrm{dist.},ab}(\boldsymbol{k}) = \frac{4\pi\sigma^2}{k} F_{ab} \int_{\sigma}^{2d+r_c} \frac{\sin(kr)}{r} \times \frac{(r_c+2d-r)^2\,(2r_c-2d+r)}{24\,d^2\,r}\,\mathrm{d}r \tag{32a}$$

and

$$\hat{p}^{\mathrm{assoc.}}_{\mathrm{ang.},ab}(\boldsymbol{k}) = \frac{4\pi\sigma^2}{k} F_{ab} \frac{(1-\cos\theta_c)^2}{4} \int_{\sigma}^{r_{12,c}} \frac{\sin(kr)}{r}\,\mathrm{d}r \tag{32b}$$

respectively, for $\boldsymbol{k} \neq \boldsymbol{0}$, such that the integrals are computed numerically. As already noticed, the weight in the Fourier space $\hat{p}_{ab}^{\mathrm{assoc.}}(\boldsymbol{k})$ is independent of the molecular distribution, thus the DFT treatment of the two potential models (distance and angular) is identical (*i.e.*, the same density profile is obtained for both formulations) as long as the weight $\hat{p}_{ab}^{\mathrm{assoc.}}(\boldsymbol{k})$ gives the same value for each potential formulation. An analogous result in the homogeneous framework of SAFT is explained by Jackson *et al.*[9] with the definition of the bonding volume (the exact relation between the association weight in the homogeneous limit and the bonding volume is given in Appendix A).

The second type of weighted function depends on the coarse-grained weight $p^{\mathrm{cg.}}$, used in Eqs. (21) and (27), for which we define the generic weighted function

$$n^{\mathrm{cg}}(\boldsymbol{r}_1) = \int f(\boldsymbol{r}_2)\, p^{\mathrm{cg}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_2 \tag{33}$$

that can be solved as

$$n^{\mathrm{cg}}(\boldsymbol{r}_1) = \mathrm{FFT}^{-1}\left[\mathrm{FFT}\left[f\right] \times \hat{p}^{\mathrm{cg}}(\boldsymbol{k})\right] \tag{34}$$

where the Fourier transform of the coarse-grained weight is given analytically, as

$$\hat{p}^{\mathrm{cg}}(\boldsymbol{k}) = \frac{\pi\sigma^3}{2}\left(\frac{\sin(k\,\sigma)}{(k\,\sigma)^3} - \frac{\cos(k\,\sigma)}{(k\,\sigma)^2}\right) \tag{35}$$

for $\boldsymbol{k} \neq \boldsymbol{0}$.

## III. RESULTS

### A. Associating fluid against a planar hard wall

To check the consistency of the association free-energy functional we consider a fluid of associating HS of radius $R = \sigma/2$ near a planar hard wall. The wall is perpendicular to the $z$-axis , and placed at $z = 0$, such that the wall potential is given by

$$V^{\text{ext.}}(z) = \begin{cases} \infty & \text{if} \quad z \leq R^+ \\ 0 & \text{if} \quad z > R^+ \end{cases} \tag{36}$$

The system has a 1D geometry, and all quantities vary along the axis $z$.

First, analytical verification can be done to check the thermodynamic coherence in the expression of a given functional. Sum rules[31] are exact relationships between microscopic properties that can be derived from a DFT model and from the macroscopic thermodynamic quantities that can be obtained from an EoS. Satisfying them in the elaboration of an EoS-based DFT model (SAFT-based in this case) corroborates that the EoS limit is well respected. The wall theorem is an application of the contact-value theorem for this system and states that the density at the contact with the wall is related to the pressure of the bulk phase (far from the wall) by[31] $\rho(R^+) = \beta P$. The pressure of an associating HS fluid can be expressed as $P = P^{\text{id.}} + P^{\text{hs}} + P^{\text{assoc.}}$, where the three terms in the right-hand side represent respectively the ideal, the HS and the association contribution to the bulk pressure. Lutsko[56] gave an analytical demonstration of the wall theorem for a van der Waals fluid. We have followed this approach for an associating HS fluid, and we have verified that the wall theorem is respected by the new functional (the detailed proof can be found in Appendix B).

In addition, numerical comparisons are done with molecular simulations at equivalent potential. An inventory of existing Monte Carlo simulations[33,34,41,43] for the associating HS fluid system can be found in the work of Camacho Vergara *et al.*[40]. The schemes of association presented in Section II A are considered with the angular cutoff $\theta_c = 27°$, and the distance cutoff $r_{12,c} = 1.05\sigma$. Molecular simulations provide density profiles $\rho(z)$ for a wide range of thermodynamic conditions, as well as the non-bonded fraction of monomers $X_0(z)$. To implement the hard wall system we consider a slit pore large enough so that the bulk fluid can be recovered far from the

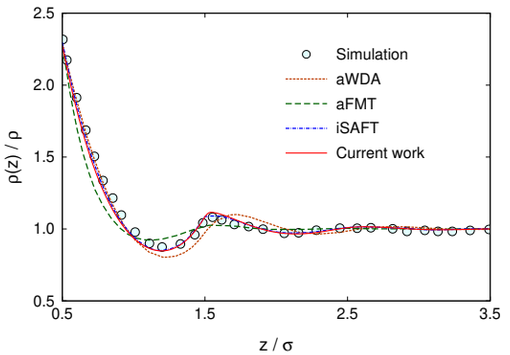
wall, such that we restrict our attention to the vicinity of the wall, where the variations in density are observed. Weighted functions and inhomogeneous excess chemical potentials of both HS and association contributions are evaluated by using FFT with a mesh width of $\sigma/125$. The density profile at equilibrium is obtained by solving the self-consistent equation (Eq. (7)) with the Picard iterative method[31]. For a pure fluid in presence of a non-associative external potential, assumptions made on the schemes imply that the non-bonded fractions of a site $a$ (*i.e.* $X_a(z)$) are all identical for any site $a$, such that the non-bonded fraction of monomers $X_0(z)$ can be related to $X_a(z)$ by the approximation[39]

$$X_0(z) = \big(X_a(z)\big)^M \tag{37}$$

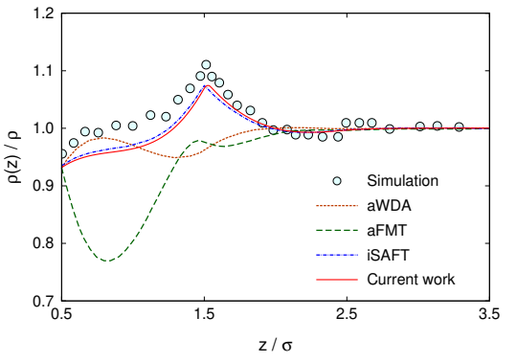
with $M$ being the number of sites for the molecule considered. The fraction $X_a(z)$ is expressed in Eq. (22) as a self-consistent equation. It is solved together with the density distribution, following the iteration steps of the Picard method such that the resulting non-bonded fraction profile is obtained for the equilibrium state. In addition to molecular simulation data, we compare the density profiles obtained in our current work to the density profiles predicted by aFMT[39], iSAFT[37] and aWDA[40]. In works presenting the aFMT as well as the aWDA, the non-bonded fraction is calculated using another method which is discussed in Appendix C 2. Thus, the non-bonded fraction profiles obtained with our functional are only compared to Monte Carlo simulations.

### *1. Density profiles*

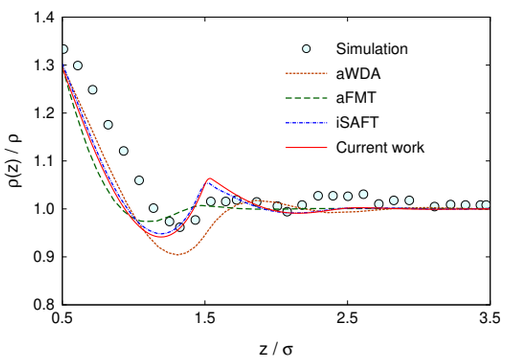
A selection of density profiles is shown in FIG. 3(a,b) for scheme 1A, in FIG. 3(c,d) for scheme 2B, and in FIG. 3(e,f) for scheme 4C. The overall shape of a density profile depends on the selected thermodynamic conditions, set by the reduced bulk density $\rho^* = \rho\sigma^3$, and the reduced association energy $\varepsilon^* = \varepsilon/(k_\mathrm{B}T)$. An increase of the density at contact with the wall is obtained when the HS contribution is predominant in the free-energy functional, as can be seen in FIG. 3(a,c,e). However, a decrease in density in the vicinity of the wall is observed when the association contribution is predominant, as shown in FIG. 3(b,d,f). Indeed, increasing the association degree of interaction competes with the repulsion effect of the hard wall, which causes a depletion of the density near the wall. When considering more sites per particle (*e.g.*, scheme 2B rather than scheme 1A) a larger amount of associating clusters within the fluid is formed, which amplifies the depletion near the wall. More details about the behaviour of an associating fluid against a planar hard wall and



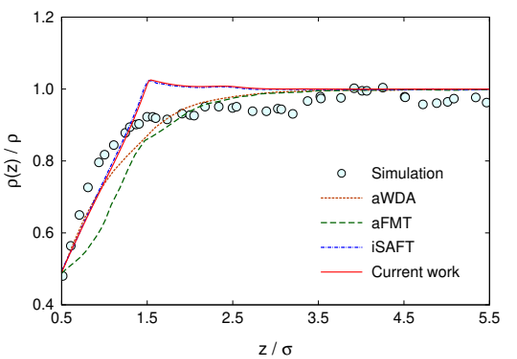

(a) Scheme 1A; $\rho^* = 0.4868$ and $\varepsilon^* = 11$



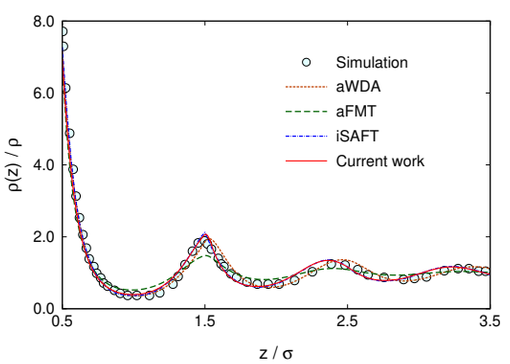

(b) Scheme 1A; $\rho^* = 0.1999$ and $\varepsilon^* = 14$



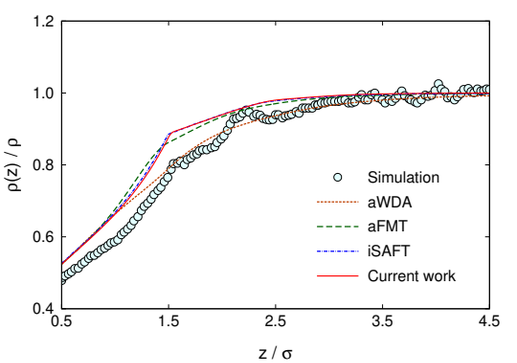

(c) Scheme 2B; $\rho^* = 0.3449$ and $\varepsilon^* = 8$




(d) Scheme 2B; $\rho^* = 0.2084$ and $\varepsilon^* = 11$



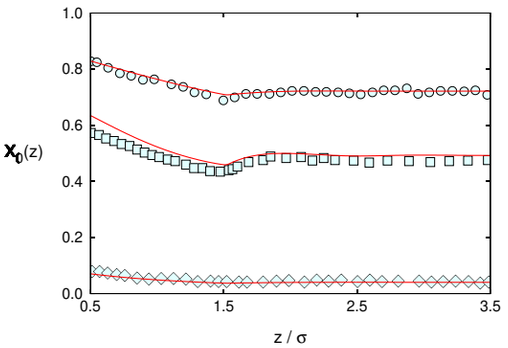

(e) Scheme 4C; $\rho^* = 0.9036$ and $\varepsilon^* = 5$



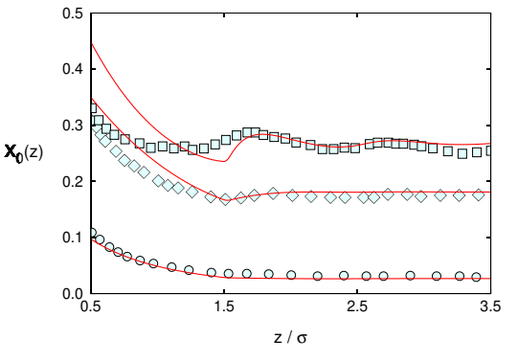

(f) Scheme 4C; $\rho^* = 0.2112$ and $\varepsilon^* = 7$

FIG. 3. Density profiles of an associating HS fluid in contact with a planar hard wall placed at $z = 0$. The results for systems described by Scheme 1A are shown in (a,b), by Scheme 2B in (c,d), and by Scheme 4C in (e,f). The profiles $\rho(z)/\rho$ are represented as a function of $z/\sigma$, where $\sigma$ is the diameter of the fluid molecules. The circles represent the simulation data[33,34,41,43]. The dotted orange line represents the aWDA calculation[40], the dashed green line represents aFMT calculation[39], the dotted-dashed blue line represents iSAFT calculation[37] and the continuous red line represents our current work.

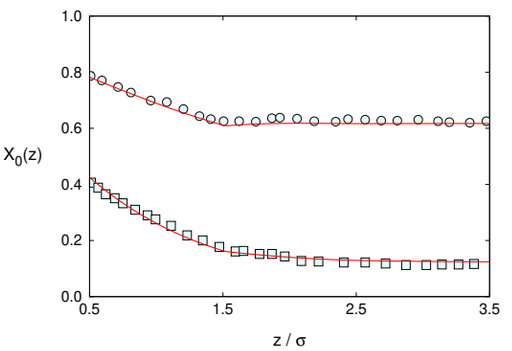
the impact of the scheme can be found in the review of Camacho Vergara *et al.*[40]. The results obtained with the free-energy functional proposed in our current work are in very good agreement with Monte Carlo simulations in all cases. As for the new functional, the density profiles are correctly captured by iSAFT, or are slightly overestimated while smoother profiles are given by both aFMT and aWDA (the peaks in density are underestimated by aFMT, while they are better reproduced by aWDA). Similar results can be observed in FIG. 3 for the new functional and iSAFT. It is not so surprising as both approaches follow a comparable guideline in the derivation : i) they use Wertheim's theory[4–7] for inhomogeneous associating fluids as a starting point (see Eqs. (11) and (12)), ii) then, an analytical expression is introduced in the PDF by extending the bulk RDF to an inhomogeneous scenario, iii) in a final step, a method is proposed for both of them to treat the inhomogeneous functions involved in the bonding probability (*i.e.* the integral over the range of association in the unbonded equation Eq. (12)). The treatment of the PDF in iSAFT and the current work differs, on the one hand, in the combination of the position-dependent RDF (Eq. (18)) and, on the other hand, in the evaluation of the contact value of the HS RDF (Eq. (20)). We could easily check the influence of the choice of combination of RDF by implementing in the current functional the geometric mean function[40] used in iSAFT. We could observe that it does not significantly change the results for the considered thermodynamic conditions (see Appendix C). As for the method of solving the integral equations (Eqs. (12) and (15)), our formalism adds no additional approximation by considering Fourier transforms to perform the calculations, as explained in section II C 4, whereas iSAFT with Segura's approximation (*cf.* Eq. (4-28) in Segura's thesis[45]) results in a 1D formulation applicable to planar geometry. As presented in appendix C 1, we managed to reformulate the original derivation of Segura[45] in a more general 3D formalism to make comparable the effects of both methods related to the inhomogeneous treatment of association. By assuming that the cutoff $r_{12,c}$ tends to the lowest limit in the range of association (*i.e.* the contact value $\sigma$), we show that solely surfacic effects are treated in an inhomogeneous manner to estimate the bonding probability in iSAFT framework. In comparison, volume effect of association is taken into account in our framework by weighting the inhomogeneous functions over the exact range of association. Segura's approximation is reliable for the cutoff used in our current work, $1.05\sigma$, which explains the close results compared to the new association function.




(a) Scheme 1A; from top to bottom: $\rho^* = 0.1984$ and $\varepsilon^* = 7$; $\rho^* = 0.4868$ and $\varepsilon^* = 7$; $\rho^* = 0.1999$ and $\varepsilon^* = 14$




(b) Scheme 2B; from top to bottom: $\rho^* = 0.7177$ and $\varepsilon^* = 6$; $\rho^* = 0.3449$ and $\varepsilon^* = 8$; $\rho^* = 0.2084$ and $\varepsilon^* = 11$




(c) Scheme 4C; from top to bottom: $\rho^* = 0.1994$ and $\varepsilon^* = 5$; and $\rho^* = 0.2112$ and $\varepsilon^* = 7$

FIG. 4. Fraction of monomers of an associating HS fluid in contact with a planar hard wall placed at $z = 0$. The results for systems described by Scheme 1A are shown in (a), by Scheme 2B in (b), and by Scheme 4C in (c). The circles, squares, and diamonds represent the simulation data[33,34,41,43]. The continuous red line represents the results obtained in the current work.

### 2. *Non-bonded fraction profiles*

The non-bonded fractions of monomers related to the system considered in the current section are shown in FIG. 4(a) for scheme 1A, in FIG. 4(b) for scheme 2B, and in FIG. 4(c) for scheme 4C. A set of thermodynamic conditions is considered for each scheme, with an increase of the association degree of interaction (from top to bottom), which tends to have a fully bonded monomer (corresponding to $X_0$ tending to 0). According to Monte Carlo simulation data[33,34,41,43], the non-bonded fraction of monomers decreases from the wall at $z = R^+$ until $z \approx 1.5\sigma$, and

remains almost constant for higher distances, for the three association schemes considered. This structure close to the wall can be easily understood as the monomers have a reduced opportunity to establish association bonds (because the wall is considered as non associative), in comparison with the monomers located far from the wall that are surrounded by other associative monomers. The results obtained with the new free-energy functional fit well the Monte Carlo simulations. Weighting the inhomogeneous functions on the exact bonding volume captures the change in the slope around $1.5\sigma$. The fraction of monomers seems overestimated for the scheme 2B for the set $\rho^* = 0.7177$ and $\varepsilon^* = 6$ (as can be seen in FIG.4(b)). However, the simulations[45] in that case can be questionable as it is the only case where no change in the slope can be seen around $1.5\sigma$, and because the fraction near the wall is similar to the fraction for the set $\rho^* = 0.3449$ and $\varepsilon^* = 8$, corresponding to a lower packing fraction.

### B. Associating fluid confined inside an anisotropic open cavity

We investigate a pure HS fluid with four associating sites (modeled by Scheme 4C), confined in a network of interconnected spherical cavities, as illustrated in FIG. 5, that mimics the shape of a Linde Type A (LTA) zeolitic nanoporous adsorbent. The open cavity was created as proposed by Bernet *et al.*[57] by considering a spherical cavity of radius $R_{\text{sph}} \approx 1.7\sigma$, truncated by a cube of edge length $a = 3.2\sigma$ placed at the center of the sphere, thus creating six spherical apertures of diameter $D_{\text{ap}} = 1.1\sigma$ along each Cartesian axis. Molecules are able to go from a cavity to another by assuming periodic boundary conditions (which is taken into account by the Fourier transforms considered for calculations). A discrete Cartesian 3D grid is created with 128 points along each axis and a mesh width of $\sigma/40$, and the origin is placed at the center of the cubic box. The WBII version of FMT[30] is used for the HS contribution, as presented in Section II B. A tensorial term[44] is added to avoid numerical divergences due to the extreme confinement of the molecules inside the cavity[57] (the tensorial contributions were not considered for the systems described in Section III A, as they can be found to be negligible, and to ensure that DFT approaches were defined at equivalent HS description). The association contribution is the one depicted in the previous Sections. The weighted functions and the chemical potential functionals of both the HS and the association contributions are evaluated with FFT. The Picard method is used to solve Eq. (7) to find the density distribution at equilibrium. For an easier interpretation, we define the bonded-fraction of monomers, $X_{0,\text{bond}}(\boldsymbol{r}) = 1 - X_0(\boldsymbol{r})$, where the non-bonded fraction of monomers $X_0(\boldsymbol{r})$

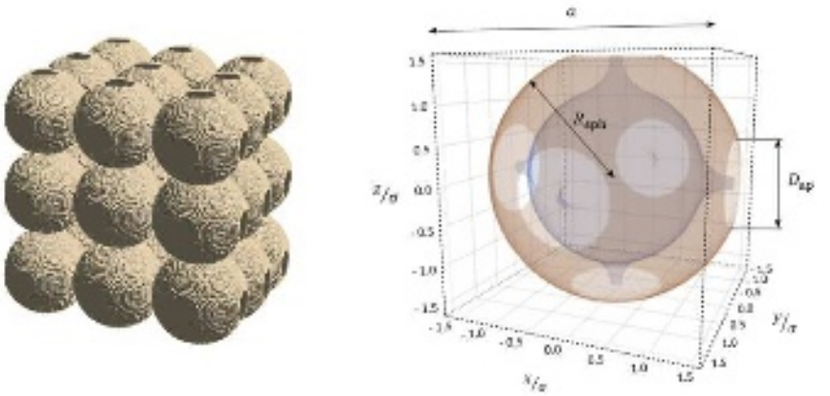
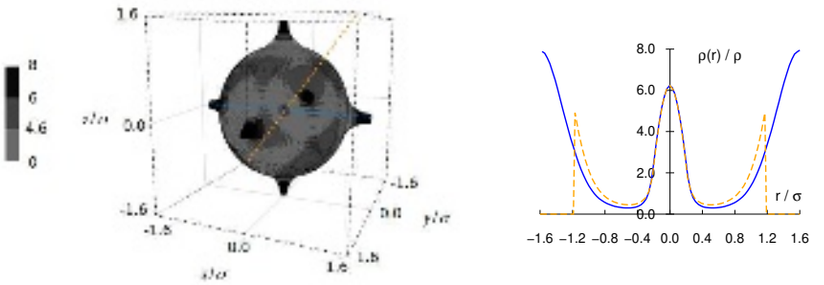


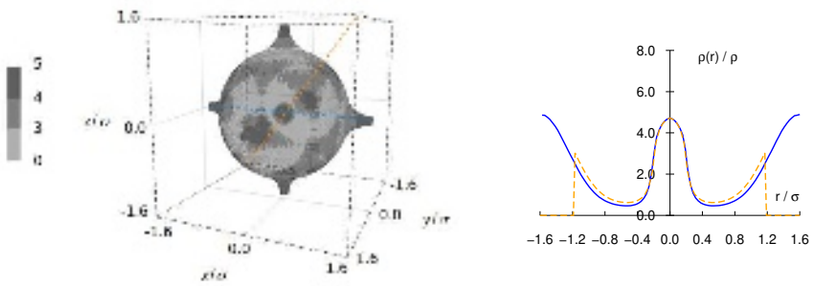

FIG. 5. Left: network of interconnected nanometer-sized hard cavities in a 3D cartesian space. Right: cubic computational box of edge length $a$ containing one cavity of the network, represented as a contour plot. A spherical cavity of radius $R_{\text{sph}} > a/2$ is placed at the center of the box, creating apertures of diameter $D_{\text{ap}}$ on the six faces of the cube. The surface of the open spherical cavity is represented in brown with a low opacity, revealing the volume accessible to the mass center of the fluid molecules, in blue.

is evaluated such as for the hard wall system.

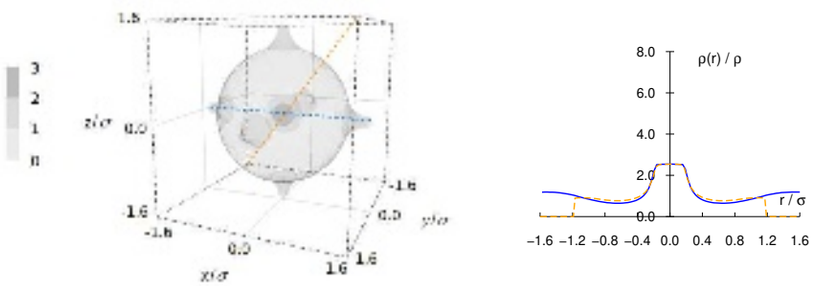
The density distributions of the associating HS inside the cavity are shown in FIG. 6 (a,b,c) for a given bulk packing fraction of $\eta = 0.3$ and different degrees of association $\varepsilon^* = 3,5,7$. A low value of association strength ($\varepsilon^* = 3$) leads to important variations of the molecular density inside the cavity (FIG. 6(a)). In that case, the preferential positions of the fluid molecules are mainly determined by the predominance of the HS contribution. The main density peaks are situated at preferential locations : in the center of the cavity, in the shell at the contact of the hard wall in eight symmetric positions distributed on the diagonals at equal distance of the channels, and inside each channel. The highest densities are respectively in the channels (eight times the bulk density), in the center (six times the bulk density), and at the eight locations at the contact of the hard cavity (five times the bulk density). The density is very close to zero between these preferential locations, except in the layer in contact with the hard cavity. The presence of association bonds leads to a smaller value of the density maxima, in comparison with the non-associating HS fluid confined in the same cavity at the same conditions[57]. When increasing the degree of association to $\varepsilon^* = 5$, the density maxima decreases in the center of the cavity, the difference in density between the eight preferential positions and the rest of the shell at the contact of the hard cavity becomes smaller (FIG. 6(b)), and the density in the channels becomes similar to the density in the center. For




(a) Density distribution for $\eta = 0.3$ and $\varepsilon^* = 3$




(b) Density distribution for $\eta = 0.3$ and $\varepsilon^* = 5$

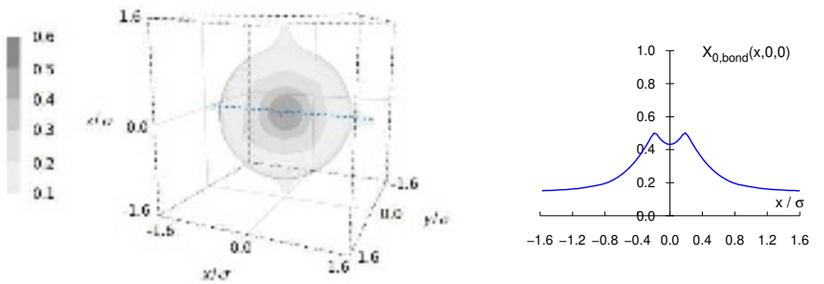



(c) Density distribution for $\eta = 0.3$ and $\varepsilon^* = 7$

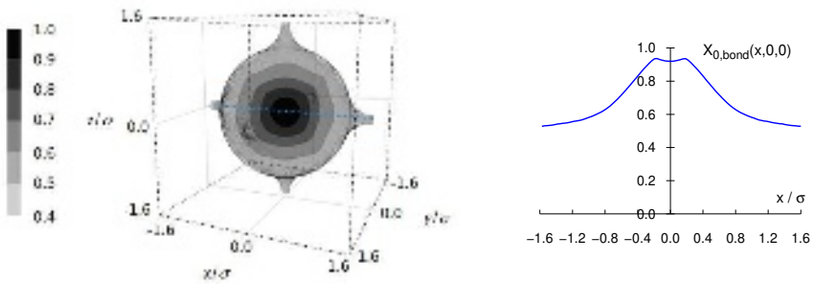
FIG. 6. Left: normalized molecular density $\rho(\boldsymbol{r})/\rho$ inside the computational box as a contour plot, the darker the higher the density. The contour values have been chosen to spotlight the preferential positions of the fluid molecules in the cavities. Right: as a quantitative supplementary information is given the density profile along the direction $(x,0,0)$ (represented with the dashed blue line in the contour plot) in continuous blue line, and the density profile along the diagonal of the box (represented with the dashed orange line in the contour plot) in dashed orange line.

$\varepsilon^* = 7$ (FIG. 6(c)), the eight preferential locations of the density peaks at the contact with the cavity disappear, leaving only a slightly higher probability in finding molecules in the center and

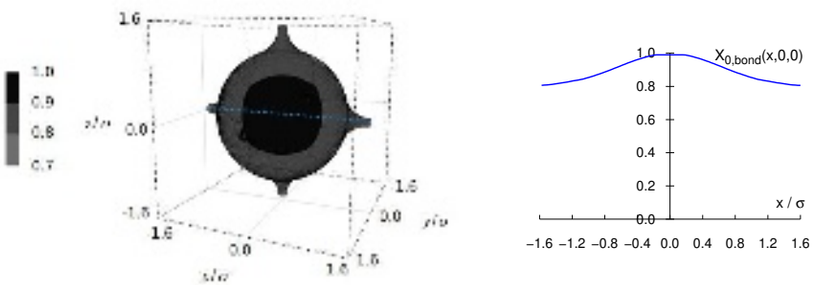
503 in the channels.




(a) Bonding fraction of monomers for $\eta = 0.3$ and $\varepsilon^* = 3$




(b) Bonding fraction of monomers for $\eta = 0.3$ and $\varepsilon^* = 5$

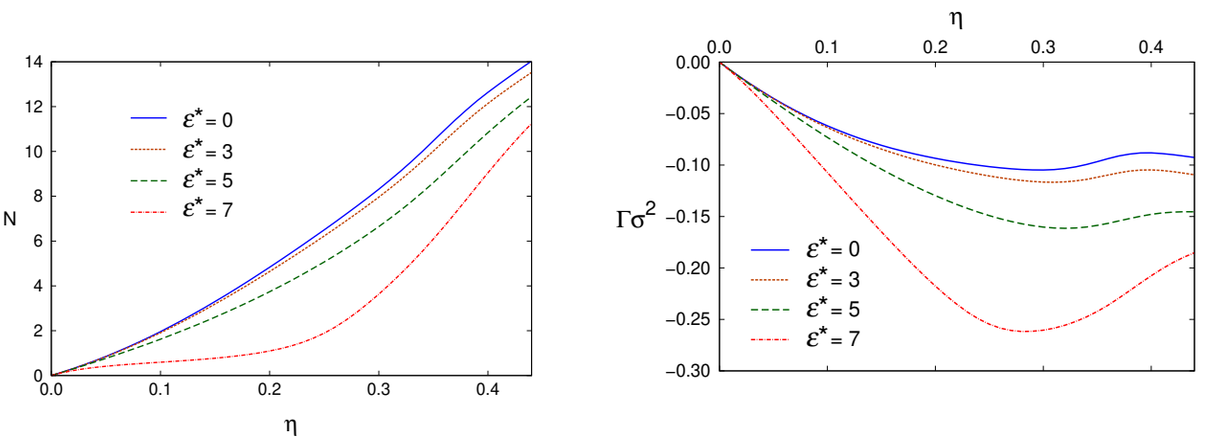



(c) Bonding fraction of monomers for $\eta = 0.3$ and $\varepsilon^* = 7$

FIG. 7. Bonded-fraction of monomer corresponding to the density profiles shown in FIG. 6. Left: the fraction is presented as a contour plot where the higher bonded fraction corresponds to the darker and opaques regions. Right: bonded fraction profile along the direction $(x,0,0)$, represented with the blue dash line in the contour plot.

504 The bonded fraction of monomers corresponding to the same set of thermodynamic conditions
505 are shown in FIG. 7(a,b,c). For a small value of association strength ($\varepsilon^* = 3$), the bonded fraction
506 of molecules is only 20% at the contact of the hard cavity and inside the channels, and it gradually

increases from this value to reach a maximum value of 50% in the center of the cavity where the molecules have the more neighbors to bond with (FIG. 7(a)). When increasing the degree of association to $\varepsilon^* = 5$, the bonded fraction of monomers is then higher than 60% and reaches more than 90% in the center (FIG. 7(b)), which are the values typically found in dense bulk associating liquids[20]. For $\varepsilon^* = 7$, fully bonded monomers are found in the major part of the cavity (FIG. 7(c)). In that case, the association energy clearly competes with the steric effect.




(a) (b)

FIG. 8. (a) Average number $N$ of associating HS inside the computational box and (b) excess adsorption $\Gamma$ as a function of the packing fraction for $\varepsilon^* = 0$ in blue (HS), $\varepsilon^* = 3$ in a dash dark-orange line, $\varepsilon^* = 5$ in a dash green line, $\varepsilon^* = 7$ dotted red line.

The average number of molecules inside the cavity, $N$, and the corresponding excess adsorption, $\Gamma$, are shown in FIG. 8. The excess adsorption is defined as

$$\Gamma = \frac{1}{A}\int \left(\rho(\boldsymbol{r}) - \rho\right) d\boldsymbol{r} = \frac{N - \rho\, V_{\text{open sphere}}}{A_{\text{open sphere}}} \tag{38}$$

where we made the choice to take into account the excluded volume in the range of integration, such that the accessible volume inside the cavity corresponds to the volume of the open sphere $V_{\text{open sphere}}$. The surface area $A$, that corresponds here to the inner envelope of surface area $A_{\text{open sphere}}$ of this volume, is introduced in Eq. (38) to return an average excess number per unit of surface. The analytical expressions of $V_{\text{open sphere}}$ and $A_{\text{open sphere}}$ are given by Bernet *et al.*[57]. The average number of molecules $N$ is evaluated numerically by integrating the molecular density $\rho(\boldsymbol{r})$ over all the computational box. For a given bulk packing fraction, the increase in the degree of association leads to a decrease in the global number of molecules that can be accommodated

inside the cavity, and thus a decrease in excess adsorption. Including association is more visible on excess adsorption for medium densities than for high densities (as steric effects are dominant); it can be interpreted as a competition between repulsive and attractive interactions.

## IV. CONCLUSIONS

A new density free-energy functional theory is proposed in this work for associating fluids. Wertheim's thermodynamics perturbation theory (TPT) is considered as the starting point of the derivation with the constraint to recover the statistical associating fluid theory (SAFT) in the bulk limit. Our framework introduces weighted functions specific to association that have a mathematical structure similar to the weighted densities of the fundamental measure theory (FMT). As these latter are convolution products, they can be easily evaluated numerically with Fourier transforms in a 3D space, thus allowing to investigate complex 3D systems. A comparison is done with existing Monte Carlo simulations and previous density functional theory formulations for a planar hard wall system (1D geometry) in order to check the consistency of the new association functional. As an example of 3D application, we investigate the extreme confinement of an associating hard sphere (HS) fluid confined in a network of 3D interconnected spherical cavities that mimics the shape of a zeolitic nanoporous adsorbent made of spherical cavities connected via channels. The impact of the degree of association on the preferential positions of the particles inside the cavity is investigated as well as the competition between on the association and the steric effect on the adsorption. For low values of association strength, the density distribution is mainly governed by the HS contribution and corresponds to a highly structured fluid with preferential positions of the molecules. While the association strength increases, the density gradients decrease and the cavity is occupied by a cluster of nearly fully bonded monomers. This leads to a lower adsorption of the associating fluid in comparison with the HS one. Although there is no available molecular simulation for this system so we could compare with, the consistency of the results we found suggests that our model is reliable to describe associating fluids confined in porous materials with complex geometry. The method we proposed to treat the inhomogeneous functions involved in the association contribution is general and can be extended to other non-associated fluid references such as square-well or soft-core potentials (*e.g.*, Lennard-Jones and Mie potentials).

## ACKNOWLEDGMENTS

The authors acknowledge the E2S UPPA hub Newpores supported by the "Investissements d'Avenir" French programme managed by ANR (ANR-16- IDEX-0002) for financial support and the PhD grant of A. Barthes. X. Liang is gratefully acknowledged for having provided all the

simulation data that are used in this work for the comparison with our results.

## DATA AVAILABILITY STATEMENT

Data underlying this article can be accessed on Zenodo at https://doi.org/10.5281/zenodo.10276406 and used under the Creative Commons Attribution license.

## REFERENCES

[1]H. Li, K. Wang, Y. Sun, C. T. Lollar, J. Li, and H.-C. Zhou, "Recent advances in gas storage and separation using metal–organic frameworks," Materials Today **21**, 108–121 (2018).

[2]Y. Ma, Z. Wang, X. Xu, and J. Wang, "Review on porous nanomaterials for adsorption and photocatalytic conversion of $co_2$," Chinese Journal of Catalysis **38**, 1956–1969 (2017).

[3]M. Ouda, F. Banat, S. W. Hasan, and G. N. Karanikolos, "Recent advances on nanotechnology-driven strategies for remediation of microplastics and nanoplastics from aqueous environments," Journal of Water Process Engineering **52**, 103543 (2023).

[4]M. S. Wertheim, "Fluids with highly directional attractive forces. I. Statistical thermodynamics." J. Stat. Phys. **35**, 19–34 (1984).

[5]M. S. Wertheim, "Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations," J. Stat. Phys. **35**, 35–37 (1984).

[6]M. S. Wertheim, "Fluids with highly directional attractive forces. III. Multiple attraction sites," J. Stat. Phys. **42**, 459–476 (1986).

[7]M. S. Wertheim, "Fluids with highly directional attractive forces. IV. Equilibrium polymerization," J. Stat. Phys. **42**, 477–492 (1986).

[8]W. Zmpitas and J. Gross, "Detailed pedagogical review and analysis of Wertheim's thermodynamic perturbation theory," Fluid Phase Equilib. **428**, 121 (2016).

[9]G. Jackson, W. G. Chapman, and K. E. Gubbins, "Phase equilibria of associating fluids spherical molecules with multiple bonding sites," Mol. Phys. **65**, 1–31 (1988).

[10]W. G. Chapman, G. Jackson, and K. E. Gubbins, "Phase equilibria of associating fluids chain molecules with multiple bonding sites," Mol. Phys. **65**, 1057 (1988).

[11]W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz, "SAFT: equation-of-state solution model for associating fluids," Fluid Phase Equilib. **52** (1989).

[12]W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz, "New reference equation of state for associating liquids," Ind. Eng. Chem. Res. **29** (1990).

[13]N. F. Carnahan and K. E. Starling, "Equation of state for nonattracting rigid spheres," J. Chem. Phys. **51**, 635–636 (1969).

[14]J. B. Schulte, P. A. Kreitzberg, C. V. Haglund, and D. Roundy, "Using fundamental measure theory to treat the correlation function of the inhomogeneous hard-sphere fluid," Phys. Rev. E **86** (2012), 10.1103/PhysRevE.86.061201.

[15]A. Gil-Villegas, A. Galindo, P. J. Whitehead, S. J. Mills, G. Jackson, and A. N. Burgess, "Statistical associating fluid theory for chain molecules with attractive potentials of variable range," J. Chem. Phys. **106**, 4168–4186 (1997).

[16]J. Gross and G. Sadowski, "Perturbed-Chain SAFT: an equation of state based on a perturbation theory for chain molecules," Ind. Eng. Chem. Res. **40**, 1244–1260 (2001).

[17]J. Gross and J. Vrabec, "An equation-of-state contribution for polar components: dipolar molecules," AIChE Journal **52**, 1194–1204 (2006).

[18]T. Lafitte, A. Apostolakou, C. Avendaño, A. Galindo, C. S. Adjiman, E. A. Müller, and G. Jackson, "Accurate statistical associating fluid theory for chain molecules formed from Mie segments," J. Chem. Phys. **139**, 154504 (2013).

[19]V. Papaioannou, T. Lafitte, C. Avendaño, C. S. Adjiman, G. Jackson, E. A. Müller, and A. Galindo, "Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments," J. Chem. Phys. **140**, 54107 (2014).

[20]S. Dufal, T. Lafitte, A. J. Haslam, A. Galindo, G. N. I. Clark, C. Vega, and G. Jackson, "The A in SAFT: developing the contribution of association to the Helmholtz free energy within a Wertheim TPT1 treatment of generic Mie fluids," Mol. Phys. **113**, 948–984 (2015).

[21]S. Dufal, T. Lafitte, A. J. Haslam, A. Galindo, G. N. I. Clark, C. Vega, and G. Jackson, "Corrigendum: The A in SAFT: developing the contribution of association to the Helmholtz free energy within a Wertheim TPT1 treatment of generic Mie fluids," Mol. Phys. **116**, 283–285 (2018).

[22]H. S. Huang and M. Radoscz, "Equation of state for small, large, polydisperse, and associating molecules," Ind. Eng. Chem. Res. **29**, 2284–2294 (1990).

[23]S. Dufal, V. Papaioannou, M. Sadeqzadeh, T. Pogiatzis, A. Chremos, C. S. Adjiman, G. Jackson, and A. Galindo, "Prediction of thermodynamic properties and phase behavior of fluids and mixtures with the SAFT-$\gamma$ Mie group-contribution equation of state," J. Chem. Eng. Data **59**,

3272–3288 (2014).

[24]S. Di Lecce, G. Lazarou, S. H. Khalit, C. S. Adjiman, G. Jackson, A. Galindo, and L. McQueen, "Modelling and prediction of the thermophysical properties of aqueous mixtures of choline geranate and geranic acid (CAGE) using SAFT-$\gamma$ Mie," RSC Adv. **9**, 38017–38031 (2019).

[25]A. J. Haslam, A. Gonzalez-Perez, S. Di Lecce, S. H. Khalit, F. A. Perdomo, S. Kournopoulos, M. Kohns, T. Lindeboom, M. Wehbe, S. A. Febra, G. Jackson, C. S. Adjiman, and A. Galindo, "Expanding the applications of the SAFT-$\gamma$ Mie group-contribution equation of state: prediction of thermodynamic properties and phase behaviour of mixtures," J. Chem. Eng. Data **65**, 5862–5890 (2020).

[26]S. A. Febra, T. Bernet, C. Mack, J. McGinty, I. I. Onyemelukwe, S. J. Urwin, J. Sefcik, J. H. ter Horst, C. S. Adjiman, G. Jackson, and A. Galindo, "Extending the SAFT-$\gamma$ Mie approach to model benzoic acid, diphenylamine, and mefenamic acid: solubility prediction and experimental measurement," Fluid Phase Equilib. **540**, 113002 (2021).

[27]R. Evans, "The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids," Adv. Phys. **28**, 143 (1979).

[28]Y. Rosenfled, "Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing," Phys. Rev. Lett. **63**, 980 (1989).

[29]Y.-X. Yu and J. Wu, "Structures of hard-sphere fluids from a modified fundamental-measure theory," J. Chem. Phys. **117** (2002).

[30]H. Hansen-Goos and R. Roth, "Density functional theory for hard-sphere mixtures: the White Bear version mark II," J. Phys.: Condens. Matter **18**, 8413 (2006).

[31]R. Roth, "Fundamental measure theory for hard-sphere mixtures: a review," J. Phys.: Condens. Matter **22**, 063102 (2010).

[32]W. G. Chapman, "Theory and Simulation of Associating Liquid Mixtures," Ph.D. Dissertation, Cornell University (1988).

[33]C. J. Segura, W. G. Chapman, and K. Shukla, "Associating fluids with four bonding sites against a hard wall: density functional theory," Mol. Phys. **90**, 759–771 (1997).

[34]C. J. Segura, E. V. Vakarin, W. G. Chapman, and M. F. Holovko, "A comparison of density functional and integral equation theories vs Monte Carlo simulations for hard sphere associating fluids near a hard wall," Mol. Phys. **108**, 4837–4848 (1998).

[35]S. Tripathi and W. G. Chapman, "Microstructure of inhomogeneous polyatomic mixtures from a density functional formalism for atomic mixtures," J. Chem. Phys. **122**, 094506 (2005).

[36]S. Jain, A. Dominik, and W. G. Chapman, "Modified interfacial statistical associating fluid theory: a perturbation density functional theory for inhomogeneous complex fluids," J. Chem. Phys. **127**, 244904 (2007).

[37]A. Bymaster and W. G. Chapman, "An iSAFT density functional theory for associating polyatomic molecules," J. Phys. Chem. **114**, 12298–12307 (2010).

[38]R. Evans, "Fundamentals of inhomogeneous fluids," Marcel-Dekker, New York (1992).

[39]Y.-X. Yu and J. Wu, "A fundamental-measure theory for inhomogeneous associating fluids," J. Chem. Phys. **116**, 7094 (2002).

[40]E. L. Camacho Vergara, G. M. Kontogeorgis, and X. Liang, "A new study of associating inhomogeneous fluids with classical density functional theory," Mol. Phys. **118**, 12298–12307 (2020).

[41]C. J. Segura and W. G. Chapman, "Associating fluids with four bonding sites against solid surfaces: Monte Carlo simulations," Mol. Phys. **86**, 415–442 (1995).

[42]P. Lurie-Gregg, J. B. Schulte, and D. Roundy, "Approach to approximating the pair distribution function of inhomogeneous hard-sphere fluids," Physical Review E **90**, 042130 (2014).

[43]A. Patrykiejew, S. Sokolowski, and D. Henderson, "The structure of associating fluids restricted by permeable walls: a density functional approach," Mol. Phys. **95**, 211–218 (1998).

[44]P. Tarazona, "Density functional for hard sphere crystals: A fundamental measure approach," Phys. Rev. Lett. **84**, 694–697 (2000).

[45]C. J. Segura, "Associating fluids near solid surfaces," Ph.D. Dissertation, Rice University (1997).

[46]M. S. Wertheim, "Fluids of dimerizing hard spheres, and fluid mixtures of hard spheres and dispheres," J. Chem. Phys. **85**, 2929 (1986).

[47]M. L. Michelsen and E. M. Hendriks, "Physical properties from association models," Fluid Phase Equilib. **180**, 165–174 (2000).

[48]A. Malijevský and S. Labík, "The bridge function for hard spheres," Mol. Phys. **60**, 663–669 (1987).

[49]J. K. Percus, "The Equilibrium Theory of Classical Fluids," (by H. L. Frisch and J. L. Lebowitz, editors, Benjamin, New York, 1964).

[50]J. K. Percus, "Approximation methods in classical statistical mechanics," Phys. Rev. Lett. **8**, 462 (1962).

[51]J. B. Schulte, P. A. Kreitzberg, C. V. Haglund, and D. Roundy, "Using fundamental measure theory to treat the correlation function of the inhomogeneous hard-sphere fluid," Phys. Rev. E

**86**, 061201 (2012).

[52]J. Gross, "A density functional theory for vapor-liquid interfaces using the PCP-SAFT equation of state," J. Chem. Phys. **131**, 204705 (2009).

[53]E. Sauer and J. Gross, "Classical density functional theory for liquid–fluid interfaces and confined systems: A functional for the perturbed-chain polar statistical associating fluid theory equation of state," Ind. Eng. Chem. Res. **56**, 4119–4135 (2017).

[54]T. Bernet, M. M. Piñeiro, F. Plantier, and C. Miqueu, "A 3D non-local density functional theory for any pore geometry," Mol. Phys. **118**, e1767308 (2020).

[55]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numercical Recipes in Fortran," Cambridge University Press (1992).

[56]J. F. Lutsko, "Recent developments in classical density functional theory," Adv. Chem. Phys. **144**, 1 (2010).

[57]T. Bernet, E. A. Müller, and G. Jackson, "A tensorial fundamental measure density functional theory for the description of adsorption in substrates of arbitrary three-dimensional geometry," J. Chem. Phys. **152**, 224701 (2020).

## Appendix A: Bulk limit

The bulk limit of the generic association weighted function defined in Eq. (29) is

$$n^{\text{assoc.}} = f \int p_{ab}^{\text{assoc.}}(r)\,\mathrm{d}\boldsymbol{r} = f\,\hat{p}_{ab}^{\text{assoc.}}(\mathbf{0}) \tag{A1}$$

where we employ the Fourier transform of $p_{ab}^{\text{assoc.}}$ for $\boldsymbol{k}=\mathbf{0}$, corresponding to the integration[9] of the distance and angular association weights (Eqs. (30a–30b)), as

$$\begin{aligned}\hat{p}_{\text{dist.},ab}^{\text{assoc.}}(\mathbf{0}) \;=\; & \frac{4\pi\sigma^2}{72d^2}\,F_{ab}\left[\ln\left(\frac{r_c+2d}{\sigma}\right)\right. \\ & \times(6r_c^3+18r_c^2d-24d^3)+(r_c+2d-\sigma) \\ & \left.\times(22d^2-5r_cd-7d\sigma-8r_c^2+r_c\sigma+\sigma^2)\right]\end{aligned} \tag{A2a}$$

and

$$\hat{p}_{\text{ang.},ab}^{\text{assoc.}}(\mathbf{0}) = 4\pi\sigma^2\,F_{ab}\,\frac{(1-\cos\theta_c)^2}{4}\,(r_{12,c}-\sigma) \tag{A2b}$$

respectively. It is convenient to define the bonding volume[9] $K_{ab}$ between sites $a$ and $b$, such that $\hat{p}_{ab}^{\text{assoc.}}(\mathbf{0}) = 4\pi K_{ab}F_{ab}$ regardless the choice of potential. We finally obtain the usual SAFT expression of the non-bonded fraction for a homogeneous fluid,

$$X_a = \frac{1}{1+\rho\,\sum_b X_b\,\Delta_{ab}} \tag{A3}$$

where $\Delta_{ab} = 4\pi\,g_\sigma^{\text{hs}}\,K_{ab}\,F_{ab}$. The bulk limit of Eq. (28) corresponds to the bulk chemical potential given by Michelsen and Hendriks[47] (under the assumption made for the contact RDF), expressed as

$$\beta\mu^{\text{assoc.}} = \sum_a \ln X_a - \frac{\rho\,\eta}{2}\,\frac{\partial \ln g_\sigma^{\text{hs}}}{\partial\eta}\,\sum_{a,b} X_a\,X_b\,\Delta_{ab} \tag{A4}$$

and the pressure is given by

$$\beta\,P^{\text{assoc.}} = -\frac{\rho^2}{2}\left(1+\eta\,\frac{\partial \ln g_\sigma^{\text{hs}}}{\partial\eta}\right)\sum_{a,b} X_a\,X_b\,\Delta_{ab} \tag{A5}$$

## Appendix B: Proof of the wall theorem for associating hard spheres

The derivation of the wall theorem to check the consistency of our new functional is detailed in this appendix. We consider a fluid of associative HSs of diameter $\sigma = 2R$ in the vicinity of a planar hard wall located at $z=0$, such that the external potential $V^{\text{ext.}}(z)$ vanishes for $z > R^+$ ( Eq. (36)).

Due to the one-dimensional geometry, the inhomogeneous quantities vary only along the axis $z$ *i.e.* perpendicular to a surface of area $A$ defined by the (infinite) planar wall. In this configuration, the Euler–Lagrange equation is given by the minimization of the grand potential $\Omega$, as

$$\frac{\delta\Omega/A}{\delta\rho(z_1)} = \frac{\delta\mathscr{F}/A}{\delta\rho(z_1)} + V^{\text{ext.}}(z_1) - \mu = 0 \tag{B1}$$

where $\mu$ is the bulk chemical potential and the derivative of the free-energy functional $\mathscr{F}$ is

$$\begin{aligned}\frac{\delta\mathscr{F}/A}{\delta\rho(z_1)} &= \frac{\delta\mathscr{F}^{\text{id.}}/A}{\delta\rho(z_1)} + \frac{\delta\mathscr{F}^{\text{hs}}/A}{\delta\rho(z_1)} + \frac{\delta\mathscr{F}^{\text{assoc.}}/A}{\delta\rho(z_1)} \\ &= \mu^{\text{id.}}(z_1) + \mu^{\text{hs}}(z_1) + \mu^{\text{assoc.}}(z_1)\end{aligned} \tag{B2}$$

For $z > R^+$

$$0 = \ln\rho(z_1) + \mu^{\text{hs}}(z_1) + \mu^{\text{assoc.}}(z_1) - \mu \tag{B3}$$

The derivation proposed by Lutsko[56] is adapted for the next steps. Both distance and angular formulations of the potential are tested, by considering a generic distance/angular associative cutoff, denoted as $r_c^{\text{assoc.}}$. Let us assume that there is a point $z_B$ such that the density is homogeneous ($\rho(z) = \rho$) for $z > z_B$, and another point, $z_b$, sufficiently far in the bulk region, so that $z_b > z_B + r_c^{\text{assoc.}}$, and $z_b > z_B + \sigma$. To bring up both the contact value of the density $\rho(R^+)$, as well as the ideal bulk pressure, Eq. (B3) is differentiated with respect to $z_1$, and multiplied by $\rho(z_1)$, as

$$0 = \frac{d\rho(z_1)}{dz_1} + \rho(z_1)\frac{d\mu^{\text{hs}}(z_1)}{dz_1} + \rho(z_1)\frac{d\mu^{\text{assoc.}}(z_1)}{dz_1} \tag{B4}$$

and integrated from $R^+$ to $z_b$, as

$$0 = \rho(z_b) - \rho(R^+) + \int_{R^+}^{z_b}\rho(z_1)\frac{d\mu^{\text{hs}}(z_1)}{dz_1}\,\mathrm{d}z_1 + \int_{R^+}^{z_b}\rho(z_1)\frac{d\mu^{\text{assoc.}}(z_1)}{dz_1}\,\mathrm{d}z_1 \tag{B5}$$

Under our assumptions, $\rho(z_b) = \rho$, such that

$$\rho(R^+) = \rho + I_{\text{hs}} + I_{\text{assoc.}} \tag{B6}$$

where $I_{\text{hs}}$ and $I_{\text{assoc.}}$ are the HS and association integrals, respectively. The bulk density is related to the ideal pressure by $\rho = \beta P^{\text{id.}}$, and $I_{\text{hs}} = \beta P^{\text{hs}}$ according to Lutsko's review[56]. By using Eq. (28) the association integral becomes

$$\begin{aligned}I_{\text{assoc.}} &= \int_{R^+}^{z_b}\rho(z_1)\frac{d\mu^{\text{assoc.}}(z_1)}{dz_1}\,\mathrm{d}z_1 \\ &= \sum_a\int_{R^+}^{z_b}\rho(z_1)\frac{d\ln X_a(z_1)}{dz_1}\,\mathrm{d}z_1 - \frac{1}{2}\sum_{a,b}\int_{R^+}^{z_b}\rho(z_1)\frac{d\bar{G}_{ab}^{\text{assoc.}}(z_1)}{dz_1}\,\mathrm{d}z_1 \\ &= \sum_a I_1 - \frac{1}{2}\sum_{a,b} I_2\end{aligned} \tag{B7}$$

Using Eq. (22), integral $I_1$ becomes

$$\begin{aligned} I_1 &= -\frac{1}{2}\sum_b \left[ \int_{R^+}^{z_b} \rho(z_1)\, X_a(z_1)\, \frac{dG_{ba}^{\text{assoc.}}(z_1)}{dz_1}\, \mathrm{d}z_1 + \int_{R^+}^{z_b} \rho(z_1)\, g_\sigma^{\text{hs}}(z_1)\, X_a(z_1)\, \frac{d\chi_{ba}(z_1)}{dz_1}\, \mathrm{d}z_1 \right. \\ &\quad \left. + \int_{R^+}^{z_b} \rho(z_1)\chi_{ba}(z_1)\, X_a(z_1)\, \frac{dg_\sigma^{\text{hs}}(z_1)}{dz_1}\, \mathrm{d}z_1 \right] \\ &= -\frac{1}{2}\sum_b \left[ I_{11} + I_{12} + I_{13} \right] \end{aligned} \tag{B8}$$

Using Eq. (27), integral $I_2$ becomes

$$I_2 = \int_{R^+}^{z_b} \rho(z_1) \int_{-\infty}^{+\infty} \rho(z_2)\, X_a(z_2)\, \frac{\partial g_\sigma^{\text{hs}}}{\partial \eta}\bigg|_{(z_2)} \chi_b(z_2)\, \frac{dp^{\text{cg}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \tag{B9}$$

such that Eq. (B7) can be expressed as

$$I_{\text{assoc.}} = -\frac{1}{2}\sum_{a,b} \left[ I_{11} + I_{12} + I_{13} + I_2 \right] \tag{B10}$$

The weight $p_{1\text{D}}^{\text{cg}}(z) = \iint p_{3\text{D}}^{\text{cg}}(x,y,z)\, \mathrm{d}x\, \mathrm{d}y$, which is the coarse-grained weight integrated over the axes $x$ and $y$ (the indices 1D are omited for in a sake of simplify) is used in Eq. (B9). In this section, we also consider the integrated association weight $p_{ab,1\text{D}}^{\text{assoc.}}(z) = \iint p_{ab,3\text{D}}^{\text{assoc.}}(x,y,z)\, \mathrm{d}x\, \mathrm{d}y$. Using Eqs. (23–24), we consider

$$\begin{aligned} \sum_{a,b} \left[ I_{11} + I_{12} \right] &= \sum_{a,b} \left[ \int_{R^+}^{z_b} \rho(z_1)\, X_a(z_1) \int_{-\infty}^{+\infty} \rho(z_2)\, X_b(z_2)\, g_\sigma^{\text{hs}}(z_2)\, \frac{dp_{ab}^{\text{assoc.}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \right. \\ &\quad \left. + \int_{R^+}^{z_b} \rho(z_1)\, g_\sigma^{\text{hs}}(z_1)\, X_a(z_1) \int_{-\infty}^{+\infty} \rho(z_2) X_b(z_2) \frac{dp_{ab}^{\text{assoc.}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \right] \end{aligned} \tag{B11}$$

by analogy with Lutsko's treatment for the van der Waals fluid[56]. Lower bounds of all the integrals in Eq. (B11) can be replaced by $z = R^+$ because $\rho(z) = 0$ for $z < R^+$. The integrals having $+\infty$ as the upper limit can be splitted by introducing the point $z_b$, so that

$$\begin{aligned} \sum_{a,b} \left[ I_{11} + I_{12} \right] &= \sum_{a,b} \left[ \int_{R^+}^{z_b} \rho(z_1)\, X_a(z_1) \int_{R^+}^{z_b} \rho(z_2)\, X_b(z_2) g_\sigma^{\text{hs}}(z_2)\, \frac{dp_{ab}^{\text{assoc.}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \right. \\ &\quad + \int_{R^+}^{z_b} \rho(z_1)\, X_a(z_1) \int_{z_b}^{\infty} \rho(z_2)\, X_b(z_2) g_\sigma^{\text{hs}}(z_2)\, \frac{dp_{ab}^{\text{assoc.}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \\ &\quad + \int_{R^+}^{z_b} \rho(z_1)\, g_\sigma^{\text{hs}}(z_1)\, X_a(z_1) \int_{R^+}^{z_b} \rho(z_2)\, X_b(z_2) \frac{dp_{ab}^{\text{assoc.}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \\ &\quad \left. + \int_{R^+}^{z_b} \rho(z_1)\, g_\sigma^{\text{hs}}(z_1)\, X_a(z_1) \int_{z_b}^{\infty} \rho(z_2)\, X_b(z_2) \frac{dp_{ab}^{\text{assoc.}}(z_{12})}{dz_1}\, \mathrm{d}z_2\, \mathrm{d}z_1 \right] \end{aligned} \tag{B12}$$

that can be rearranged as

$$\begin{aligned}\sum_{a,b}[I_{11}+I_{12}] \;=\; & \sum_{a,b}\Bigg[\int_{R^+}^{z_b}\int_{R^+}^{z_b}\rho(z_1)\,\rho(z_2)\,X_a(z_1)\,X_b(z_2)\left(g_\sigma^{\mathrm{hs}}(z_1)+g_\sigma^{\mathrm{hs}}(z_2)\right)\frac{dp_{ab}^{\mathrm{assoc.}}(z_{12})}{dz_1}\,\mathrm{d}z_2\,\mathrm{d}z_1\Bigg] \\ & +\sum_{a,b}\Bigg[\int_{R^+}^{z_b}\rho(z_1)\,X_a(z_1)\int_{z_b}^{+\infty}\rho(z_2)\,X_b(z_2)g_\sigma^{\mathrm{hs}}(z_2)\,\frac{dp_{ab}^{\mathrm{assoc.}}(z_{12})}{dz_1}\,\mathrm{d}z_2\,\mathrm{d}z_1 \\ & +\int_{R^+}^{z_b}\rho(z_1)\,g_\sigma^{\mathrm{hs}}(z_1)\,X_a(z_1)\int_{z_b}^{+\infty}\rho(z_2)\,X_b(z_2)\frac{dp_{ab}^{\mathrm{assoc.}}(z_{12})}{dz_1}\,\mathrm{d}z_2\,\mathrm{d}z_1\Bigg]\end{aligned} \tag{B13}$$

The first sum on the right-hand side can be cancelled since it is odd by permutation of $z_1$ into $z_2$. The last sum can also be rearranged since $\rho(z_2)$, $X_b(z_2)$ and $g_\sigma^{\mathrm{hs}}(z_2)$ become constant and equal to their bulk value in the region $z_2 > z_b$, such that

$$\begin{aligned}\sum_{a,b}[I_{11}+I_{12}] \;=\; & \rho\sum_{a,b}X_b\Bigg[g_\sigma^{\mathrm{hs}}\int_{R^+}^{z_b}\rho(z_1)\,X_a(z_1)\int_{z_b}^{+\infty}\frac{dp_{ab}^{\mathrm{assoc.}}(z_{12})}{dz_1}\,\mathrm{d}z_2\,\mathrm{d}z_1 \\ & +\int_{R^+}^{z_b}\rho(z_1)\,g_\sigma^{\mathrm{hs}}(z_1)\,X_a(z_1)\int_{z_b}^{+\infty}\frac{dp_{ab}^{\mathrm{assoc.}}(z_{12})}{dz_1}\,\mathrm{d}z_2\,\mathrm{d}z_1\Bigg] \\ = \; & \rho\sum_{a,b}X_b\Bigg[g_\sigma^{\mathrm{hs}}\int_{R^+}^{z_b}\rho(z_1)\,X_a(z_1)p_{ab}^{\mathrm{assoc.}}(z_1-z_b)\,\mathrm{d}z_1 \\ & +\int_{R^+}^{z_b}\rho(z_1)\,g_\sigma^{\mathrm{hs}}(z_1)\,X_a(z_1)p_{ab}^{\mathrm{assoc.}}(z_1-z_b)\,\mathrm{d}z_1\Bigg]\end{aligned} \tag{B14}$$

Making use of the associative cutoff

$$\begin{aligned}\sum_{a,b}[I_{11}+I_{12}] \;=\; & \rho\sum_{a,b}X_b\Bigg[g_\sigma^{\mathrm{hs}}\int_{z_b-r_c^{assoc.}}^{z_b}\rho(z_1)\,X_a(z_1)p_{ab}^{\mathrm{assoc.}}(z_1-z_b)\,\mathrm{d}z_1 \\ & +\int_{z_b-r_c^{assoc.}}^{z_b}\rho(z_1)\,g_\sigma^{\mathrm{hs}}(z_1)\,X_a(z_1)p_{ab}^{\mathrm{assoc.}}(z_1-z_b)\,\mathrm{d}z_1\Bigg]\end{aligned} \tag{B15}$$

Under the considered assumptions, $\rho(z_1)$, $X_a(z_1)$ and $g_\sigma^{\mathrm{hs}}(z_1)$ become constant and equal to their bulk value in the region $z_1 > z_b$ , such that

$$\begin{aligned}\sum_{a,b}[I_{11}+I_{12}] \;&= 2\rho^2\,g_\sigma^{\mathrm{hs}}\sum_{a,b}X_a\,X_b\int_0^{r_c^{\mathrm{assoc.}}}p_{ab}^{\mathrm{assoc.}}(z_1)\,\mathrm{d}z_1 \\ &= \rho^2\sum_{a,b}X_a\,X_b\,\Delta_{ab}\end{aligned} \tag{B16}$$

by using $4\pi\,K_{ab}\,F_{ab} = \int_{-r_c^{\mathrm{assoc.}}}^{r_c^{\mathrm{assoc.}}} p_{ab}^{\mathrm{assoc.}}(z_1)\,\mathrm{d}z_1$. The two remaining integrals in Eq. (B10) can be simplified in a similar way, and give

$$\begin{aligned}\sum_{a,b}[I_{13}+I_2] \;=\; & \sum_{a,b}\Bigg[\int_{0^+}^{z_b}\rho(z_1)\chi_{ba}(z_1)X_a(z_1)\,\frac{dg_\sigma^{\mathrm{hs}}(z_1)}{dz_1}\,\mathrm{d}z_1 \\ & +\int_{R^+}^{z_b}\rho(z_1)\int_{-\infty}^{+\infty}\rho(z_2)\,X_a(z_2)\,\left.\frac{\partial g_\sigma^{\mathrm{hs}}}{\partial\eta}\right|_{(z_2)}\chi_{ba}(z_2)\,\frac{dp^{\mathrm{cg}}(z_{12})}{dz_1}\,\mathrm{d}z_2\,\mathrm{d}z_1\Bigg]\end{aligned} \tag{B17}$$

where we used

$$\begin{aligned}\frac{dg_\sigma^{\rm hs}(z_1)}{dz_1} &= \left.\frac{\partial g_\sigma^{\rm hs}}{\partial \eta}\right|_{(z_1)} \frac{\partial \eta}{\partial z_1} \\ &= \left.\frac{\partial g_\sigma^{\rm hs}}{\partial \eta}\right|_{(z_1)} \int_{-\infty}^{+\infty} \rho(z_2)\,\frac{dp^{\rm cg}(z_{12})}{dz_1}\,{\rm d}z_2\end{aligned} \tag{B18}$$

With the same methodology used in Eqs. (B12–B13), we obtain

$$\begin{aligned}\sum_{a,b}[I_{13}+I_2] \;=\; &\sum_{a,b}\left[\int_{R^+}^{z_b}\int_{R^+}^{z_b}\rho(z_1)\,\rho(z_2)\left(X_a(z_1)\,\chi_{ba}(z_1)\;\left.\frac{\partial g_\sigma^{\rm hs}}{\partial\eta}\right|_{(z_1)} + X_a(z_2)\,\chi_{ba}(z_2)\;\left.\frac{\partial g_\sigma^{\rm hs}}{\partial\eta}\right|_{(z_2)}\right){\rm d}z_2\,{\rm d}z_1\right] \\ &+\sum_{a,b}\left[\int_{R^+}^{z_b}\rho(z_1)\chi_{ba}(z_1)X_a(z_1)\;\left.\frac{\partial g_\sigma^{\rm hs}}{\partial\eta}\right|_{(z_1)}\int_{z_b}^{+\infty}\rho(z_2)\,\frac{dp^{\rm cg}(z_{12})}{dz_1}\,{\rm d}z_2\,{\rm d}z_1\right. \\ &\left.+\int_{R^+}^{z_b}\rho(z_1)\int_{z_b}^{+\infty}\rho(z_2)\,X_a(z_2)\;\left.\frac{\partial g_\sigma^{\rm hs}}{\partial\eta}\right|_{(z_2)}\chi_{ba}(z_2)\,\frac{dp^{\rm cg}(z_{12})}{dz_1}\,{\rm d}z_2\,{\rm d}z_1\right]\end{aligned} \tag{B19}$$

Performing a derivation similar to the one done in Eqs. (B14–B16) with the CG cutoff $\sigma$ allows to obtain

$$\begin{aligned}\sum_{a,b}[I_{13}+I_2] &= 2\rho^2\,\frac{\partial g_\sigma^{\rm hs}}{\partial\eta}\int_0^\sigma p^{\rm cg}(z_1)\,{\rm d}z_1\sum_{a,b}X_a\,\chi_{ba} \\ &= \rho^2\,\eta\,\frac{\partial \ln g_\sigma^{\rm hs}}{\partial\eta}\sum_{a,b}X_a\,X_b\,\Delta_{ab}\end{aligned} \tag{B20}$$

and Eq. (B10) becomes

$$\begin{aligned}I_{\rm assoc.} &= -\frac{\rho^2}{2}\left(1+\eta\frac{\partial \ln g_\sigma^{\rm hs}}{\partial\eta}\right)\sum_{a,b}X_a\,X_b\,\Delta_{ab} \\ &= \beta\,P^{\rm assoc.}\end{aligned} \tag{B21}$$

such that Eq. (B6) reduces to $\rho(R^+) = \beta P$ and the wall theorem is verified.

## Appendix C: Inhomogeneous treatment of the bonding fraction

The probability of finding an elegible bonding partner, for a site $a$ on a given molecule at position $\boldsymbol{r}_1$, is expressed through a complex pair correlation integral in the non-bonded fraction relation, that we recall here:

$$X_a(\boldsymbol{r}_1) = \frac{1}{1+\int\rho(\boldsymbol{r}_2)\,g^{\rm hs}(\boldsymbol{r}_1,\boldsymbol{r}_2)\sum_b X_b(\boldsymbol{r}_2)\,\bar{F}_{ab}(r_{12})\,{\rm d}\boldsymbol{r}_2} \tag{C1}$$

In the original derivation proposed by Wertheim[4–7], this bonding probability depends on both the position of the associating molecules and the orientation of theirs respective sites. An average over sites orientations is convenient so that the formalism is simplified to a position-dependant density functional theory. The range of association around a given molecule in a spherical envelope, represented in FIG. 9, is defined by the orientation-averaged Mayer function $\bar{F}_{ab}(r_{12})$. For the angular potential formulation, this range corresponds to an interval from the contact value between two molecules (*i.e.*, the diameter $\sigma$) to the cutoff $r_{12,c}$, such that the bonding is restricted to a dimer formation[9] (*i.e.*, $\sigma < r_{12,c} \leq \sigma/[2\sin(\theta_c)]$). To estimate the bonding probability in an inhomogeneous fluid, one requires the values of the local functions (*e.g.*, local density) in the defined region. Segura *et al.*[41] was the first to drive Wertheim's unbonded fraction relation to a tractable equation for inhomogeneous fluid. It was later enhanced by Chapman and coworkers[37] in iSAFT framework.

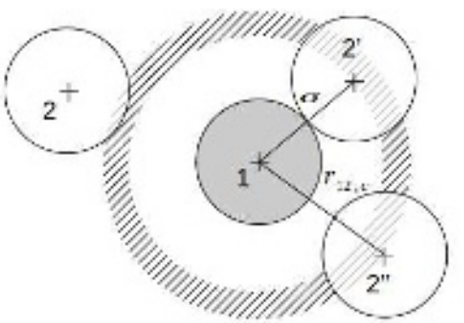



FIG. 9. Schematic representation of the range of association between two molecules placed at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, respectively. The lower limit of the range is represented with molecule 2', and the upper limit of the range is represented with molecule 2". A non-bonded case is represented with molecule 2.

A reformulation of the original derivation[45], under the same assumptions, is given in Section C 1 of this appendix to make clear the improvements provided by our current work concerning the inhomogeneous treatment of association. In Section C 2, we discuss the method proposed by Yu and Wu[39] (and employed by Camacho Vergara *et al.*[40]) to calculate the unbonded fraction profiles.

### 1. Surface and volume effect of association

We consider here the case of an associating fluid with four sites within the scheme 4C. The

expression of the non-bounded fraction obtained by Segura *et al.*[41,45] is given as a 1D formulation as

$$X_a^{\text{Segura}}(z_1) = \frac{1}{1 + \dfrac{4\,\pi\,F_{aa}\,K_{aa}\,g^{\text{hs}}(\sigma\,;\,\rho_{\text{bulk}})}{\sigma} \displaystyle\int_{z_1-\sigma}^{z_1+\sigma} \rho(z_2)\,X_a(z_2)\,\mathrm{d}z_2} \tag{C2}$$

Two assumptions are made to obtain Eq. (C2) from Eq. (C1). The first assumption deals with the pair correlation function (and corresponds to the approximation that Jackson *et al.*[9] employed in SAFT for homogeneous fluids), treated as $g^{\text{hs}}(\boldsymbol{r}_1,\boldsymbol{r}_2) \approx \dfrac{\sigma^2}{{r_{12}}^2} g^{\text{hs}}(\sigma\,;\,\rho_{\text{bulk}})$, where $\rho_{\text{bulk}}$ is the bulk density for the considered chemical potential $\mu$ (introduced in Eq. (5)), and $g^{\text{hs}}(\sigma\,;\,\rho_{\text{bulk}})$ is the contact-value of the pair distribution function of the homogeneous hard-sphere (HS) fluid for $\rho_{\text{bulk}}$. By using these assumptions in Eq. (C1), this latter becomes

$$X_a(\boldsymbol{r}_1) = \frac{1}{1 + 2\,g^{\text{hs}}(\sigma\,;\,\rho_{\text{bulk}})\,F_{aa}\,K_{aa} \displaystyle\int \rho(\boldsymbol{r}_2)\,X_a(\boldsymbol{r}_2)\,\frac{1}{{r_{12}}^2}\,\frac{\Theta(r_{12,c}-r_{12})-\Theta(\sigma-r_{12})}{r_{12,c}-\sigma}\,\mathrm{d}\boldsymbol{r}_2} \tag{C3}$$

where we express the average Mayer function $\bar{F}_{aa}(r_{12})$ with the weights defined in Eq. (14b) (for the angular formulation of the potential) and introduce the bonding volume $K_{aa}$, defined by Segura *et al.*[41,45] as $K_{aa} = \dfrac{(1-\cos\theta_c)^2}{4}\sigma^2\,(r_{12,c}-\sigma)$ (which in turn introduces the factor $(r_{12,c}-\sigma)$ in the denominator).

The second approximation consists in simplifying the integral in Eq. (C3) by considering that the bonding range is small. If $r_{12,c}$ tends to $\sigma^+$, one can note that

$$\lim_{r_{12,c}\to\sigma^+} \frac{\Theta(r_{12,c}-r_{12})-\Theta(\sigma-r_{12})}{r_{12,c}-\sigma} = \delta(\sigma-r_{12}) \tag{C4}$$

where $\delta(\sigma-r_{12})$ is a Dirac distribution. As a consequence, the term $r_{12}^2$ in Eq. (C3) can be replaced by $\sigma^2$, such that the non-bonded fraction can be expressed as

$$X_a^{\text{3D,Segura}}(\boldsymbol{r}_1) = \frac{1}{1 + \dfrac{2\,g^{\text{hs}}(\sigma\,;\,\rho_{\text{bulk}})\,F_{aa}\,K_{aa}}{\sigma^2} \displaystyle\int \rho(\boldsymbol{r}_2)\,X_a(\boldsymbol{r}_2)\,\delta(\sigma-r_{12})\,\mathrm{d}\boldsymbol{r}_2} \tag{C5}$$

This expression is equivalent to the 1D formulation (Eq. (C2)) in a more general 3D form, where the local functions $\rho(\boldsymbol{r})$ and $X_a(\boldsymbol{r})$ are weighted (*i.e.*, averaged) on the surface of a sphere of radius $\sigma$, that corresponds to the accessible surface area surrounding the molecule at position $\boldsymbol{r}_1$. The approximation used by Jackson *et al.*[9] to estimate the pair correlation beyond the contact value vanishes by using the Dirac distribution. Bymaster *et al.*[37] improved the initial formulation of Segura, by replacing the homogeneous pair correlation function with an inhomogeneous scenario

by using position-reference densities, but retained the second approximation (using of the Dirac distribution defined in Eq. (C4)) within a 1D-equivalent formulation[37]).

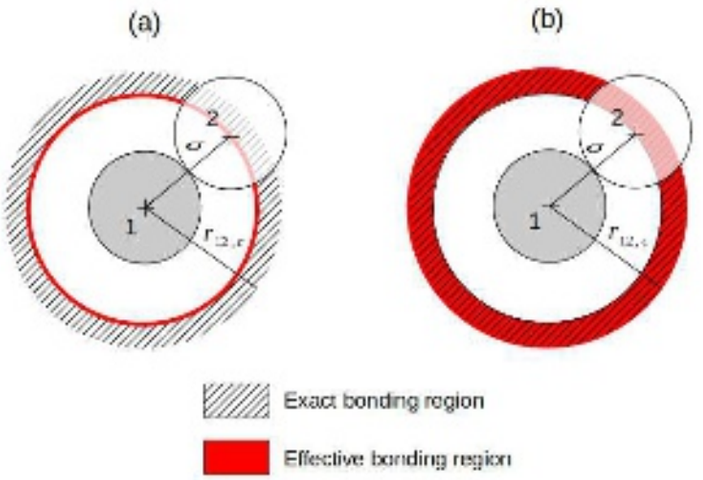



FIG. 10. Schematic representation of the exact and effective regions of association between molecules 1 and 2, measured in $\boldsymbol{r}_1$ (the region of association occurs in the estimation of the unbonded fraction $X_a(\boldsymbol{r}_1)$ for molecule 1). The exact bonding region, represented as a hatched zone, is delimited radially by the contact distance between molecule 1 and molecule 2, *i.e.*, the diameter $\sigma$, and by the cutoff $r_{12,c}$. Case (a) refers to the approximation employed by Segura[41,45] and iSAFT[37] models, where the inhomogeneous functions are weigthed on the accessible surface only, and where volume effects are neglected for association in an inhomogeneous manner. Case (b) refers to the new treatment proposed in our current work, where the inhomogeneous functions are weighted over the exact bonding volume of association.

With this second approximation, the treatment of inhomogeneous functions involved in the bonding volume is restricted to surface effects (due to the Dirac distribution), as it is illustrated in FIG. 10. The effects of the approximation are not negligible when the cutoff $r_{12,c}$ is no longer similar to the contact value $\sigma^+$. We compared the formulation proposed in the current work (normalized by the bonding volume and denoted as $\bar{n}^{\text{assoc.}}(\boldsymbol{r}_1)$ here), which depends on the cutoff $r_{12,c}$,

$$\begin{aligned}\bar{n}^{\text{assoc.}}(\boldsymbol{r}_1) &= \frac{\displaystyle\int f(\boldsymbol{r}_2)\frac{1}{{r_{12}}^2}\Big(\Theta(r_{12,c}-r_{12})-\Theta(\sigma-r_{12})\Big)\,\mathrm{d}\boldsymbol{r}_2}{\displaystyle\int \frac{1}{{r_{12}}^2}\Big(\Theta(r_{12,c}-r_{12})-\Theta(\sigma-r_{12})\Big)\,\mathrm{d}\boldsymbol{r}_2} \\ &= \frac{1}{4\pi}\int f(\boldsymbol{r}_2)\frac{1}{{r_{12}}^2}\,\frac{\Theta(r_{12,c}-r_{12})-\Theta(\sigma-r_{12})}{r_{12,c}-\sigma}\,\mathrm{d}\boldsymbol{r}_2\end{aligned} \tag{C6}$$

with the approximate function obtained by considering $r_{12,c}\to\sigma^+$ (also normalized by the bond-

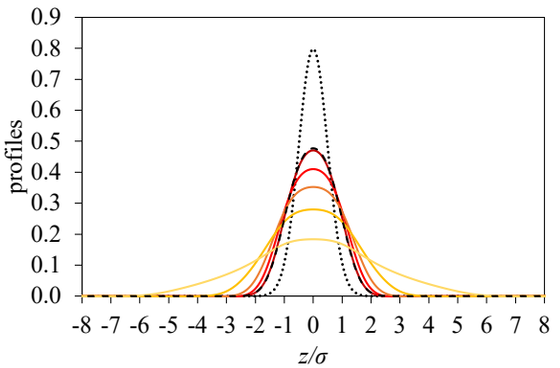



FIG. 11. Illustration of the cutoff influence in the association weight. A Gaussian profile $f(z)$ of reference is represented by a dotted curve. The corresponding weighted function $\bar{n}^{\text{Segura}}(z)$ defined from Segura's treatment[45] is represented by a dashed curve. The corresponding weighted function $\bar{n}^{\text{assoc.}}(z)$ introduced in our current work is represented by continuous curves for several cutoffs: $1.05\sigma$ (dark red), $1.46\sigma$ (red), $1.93\sigma$ (dark orange), $2.88\sigma$ (orange), and $5.74\sigma$ (yellow).

ing volume, and denoted as $\bar{n}^{\text{Segura}}(\boldsymbol{r_1})$ here), that is independent of the cutoff,

$$\bar{n}^{\text{Segura}}(\boldsymbol{r_1}) = \frac{\int f(\boldsymbol{r}_2)\,\delta(\sigma - r_{12})\,\mathrm{d}\boldsymbol{r}_2}{\int \delta(\sigma - r_{12})\,\mathrm{d}\boldsymbol{r}_2} = \frac{1}{4\pi\sigma^2}\int f(\boldsymbol{r}_2)\,\delta(\sigma - r_{12})\,\mathrm{d}\boldsymbol{r}_2 \tag{C7}$$

For a 1D Gaussian profile centered in zero,

$$f(z) = \frac{1}{\sqrt{2\pi}\Delta}\exp\left(-\frac{z^2}{2\Delta^2}\right) \tag{C8}$$

we show the results in fig. 11, for $\Delta = \sigma/2$. For the cutoff employed in our current work, $1.05\sigma$, the difference between $\bar{n}^{\text{Segura}}(z)$ and $\bar{n}^{\text{assoc.}}(z)$ is negligible, explaining the close results between the new association functional and iSAFT. The larger the cutoff is, the more spread out is the function $\bar{n}^{\text{assoc.}}(z)$. In addition, we consider the following cutoff radius: $1.46\sigma$, $1.93\sigma$, $2.88\sigma$, and $5.74\sigma$, that correspond to the maximum possible cutoffs (to respect the dimerization condition) for angles $\theta_c$ of 20°, 15°, 10°, and 5°, respectively.

### 2. Two methods for the calculation of the non-bonded fraction

Yu and Wu[39] developed a DFT model for associating fluids, denoted as aFMT here, by introduc-

ing the weighted densities of the fundamental-measure theory (FMT), $n_\alpha(\boldsymbol{r})$ in their framework (where $\alpha$ denotes the nature of the weighting[30,44]). The non-bonded fraction of an associating fluid described by scheme 4C with the aFMT approach is expressed

$$X_a^{\mathrm{aFMT}}(\boldsymbol{r}) = \frac{1}{1+2\,n_0(\boldsymbol{r})\zeta(\boldsymbol{r})X_a(\boldsymbol{r})\,\Delta_{aa}(\boldsymbol{r})} \tag{C9}$$

where $\zeta(\boldsymbol{r})$ is a function of scalar and vectorial FMT's weighted densities. The association strength function $\Delta_{aa}(\boldsymbol{r})$ is given by

$$\Delta_{aa}(\boldsymbol{r}) = 4\pi K_{aa}\,F_{aa}\,g(\sigma, n_\alpha) \tag{C10}$$

where $K_{aa}$ is the constant bonding volume, defined above, $F_{aa}$ is the Mayer function, and $g(\sigma, n_\alpha)$ is the pair correlation between two HS molecules at contact, expressed as a function of the FMT's weighted densities $n_\alpha(\boldsymbol{r})$. The inhomogeneous functions used to estimate the bonding fraction of the monomer correspond to averages over the surface area and volume of the monomer itself through the FMT's weighted densities. These contributions are then multiplied by the bonding volume. Another formulation of the non-bonding fraction is proposed by Yu and Wu[39], consisting in applying Segura's expression (*i.e.*, Eq. (C2)) to the density profile obtained with the aFMT approach. The resulting non-bonded fraction is denoted as $X_a^{\mathrm{aFMT-Segura}}(\boldsymbol{r})$.

We show in FIG.12 the 1D non-bonded fraction profiles, $X_0(z)$ (calculated with Eq. (37) from the fractions $X_a(z)$), of an associating HS fluid against a planar hard wall for two sets of thermodynamic conditions, characterized by $\rho^* = 0.1994$ and $\varepsilon^* = 5$, and by $\rho^* = 0.2112$ and $\varepsilon^* = 7$, respectively. We compare the prediction obtained with the aFMT approach, the aFMT-Segura approach, the new functional developed in the current work, and molecular simulations at equivalent potential. A change of slope is observed in molecular simulations at $z \approx 1.5\sigma$, which is correctly predicted by the new functional and by the aFMT-Segura approach. A change of slope at $z \approx \sigma$ is however observed with the aFMT approach, which does not correspond to the trend of the molecular simulations. These differences can be explained by the weighted functions used in the approaches considered here: for a hard wall placed at $z = 0$ (*i.e.*, for an accessible region of $z > \sigma/2$ for the center of fluid particles) and a weight characterized by a radius $\mathscr{R}$, a change in slope can be observed at $z \approx \sigma/2 + \mathscr{R}$. In particular, the fraction $X_0^{\mathrm{aFMT}}(z)$ (Eq. (C9)) depends on weights of radius $\sigma/2$ only. By contrast, an effective weight using the contact distance $\sigma$ is introduced by the aFMT-Segura approximation (Eq. (C2)) to calculate $X_0^{\mathrm{aFMT-Segura}}(z)$. The association contribution in the new free-energy functional depends on weights employing the contact

distance $\sigma$, and the bonding volume of association from $\sigma$ to $r_{12,c} = 1.05\sigma$ (Eq. (22)), which explains that the change in slope occurs at $z \approx 1.5\sigma$.

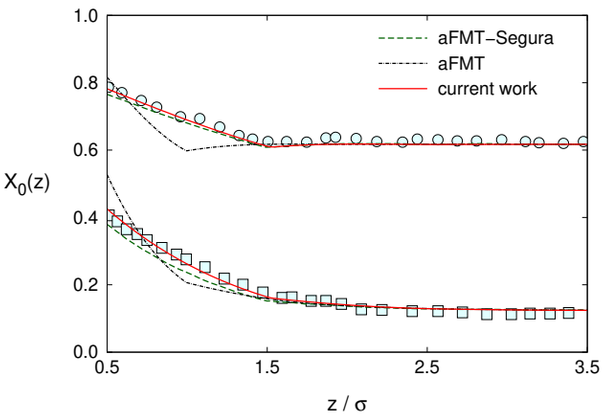
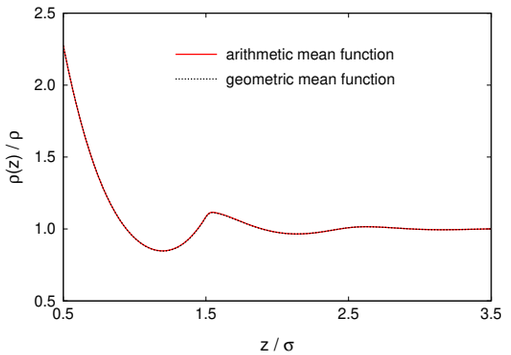


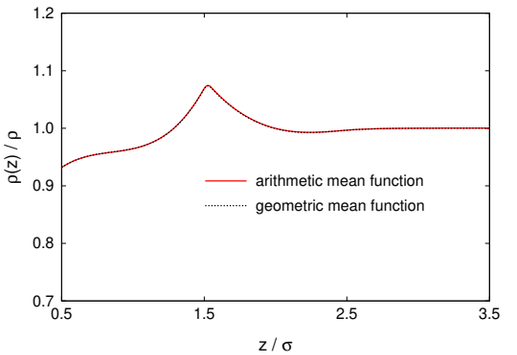

FIG. 12. Fraction of non-bonded monomers of an associated fluid with four associating sites (*i.e.*, scheme 4C) in contact with a planar hard wall placed at $z = 0$. From top to bottom: $\rho^* = 0.1994$ and $\varepsilon^* = 5$; and $\rho^* = 0.2112$ and $\varepsilon^* = 7$. The circles and squares represent the simulation data[33,34,41,43]. The dashed green line represents the aFMT-Segura calculation (Eq. (C2)). The dotted-dashed black line represents aFMT result with the unbonded fraction expression obtained by Yu and Wu[39] (Eq. (C9)). The continuous red line represents the results obtained in the current work (Eq. (22)).

## Appendix D: Two mean functions to approximate the pair distribution function (PDF)

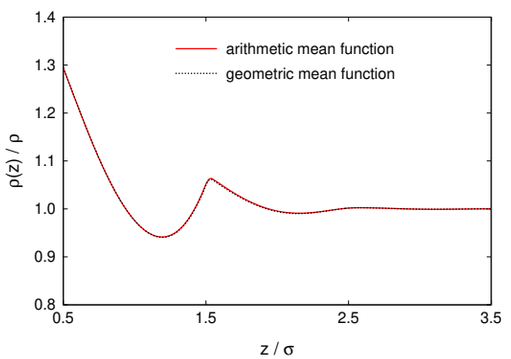
As discussed by Lurie-Gregg *et al.*[42], a separation of the variables $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in the PDF, $g^{\mathrm{hs}}(\boldsymbol{r}_1, \boldsymbol{r}_2)$, is convenient so that the pair correlation integrals reduce to one-centers convolution products. To derive our association DFT model that we detailed in Section II C, we employed the arithmetic mean function by analogy with the contact-value approach proposed by Lurie-Gregg *et al.*[42]. However one could also use a geometric mean function, $g^{\mathrm{hs}}(\boldsymbol{r}_1, \boldsymbol{r}_2) \approx \left[ g^{\mathrm{hs}}(r_{12}\,;\,\eta_\sigma(\boldsymbol{r}_1))\, g^{\mathrm{hs}}(r_{12}\,;\,\eta_\sigma(\boldsymbol{r}_2)) \right]^{1/2}$, as it has been done in iSAFT calculation[37], by keeping the remaining part of the derivation unchanged. We compare in FIG.13 both formulations applied to the hard wall system that we presented in Section II C for the same association scheme and identical thermodynamic conditions, and we do not obtain significant changes in the results.



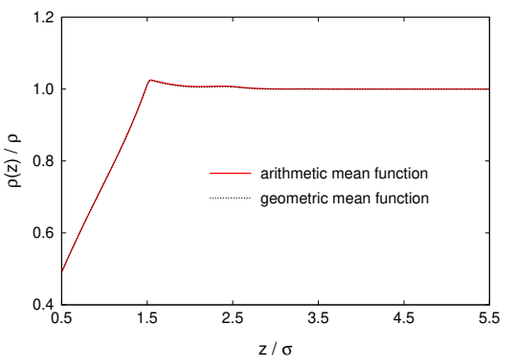

(a) Scheme 1A; $\rho^* = 0.4868$ and $\varepsilon^* = 11$



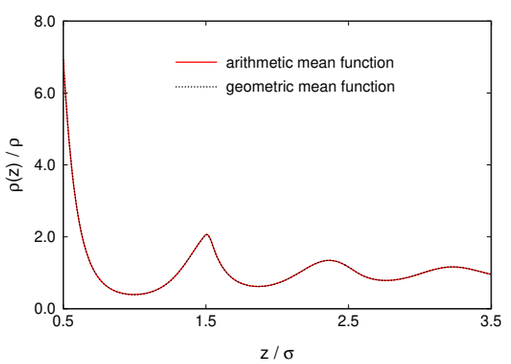

(b) Scheme 1A; $\rho^* = 0.1999$ and $\varepsilon^* = 14$



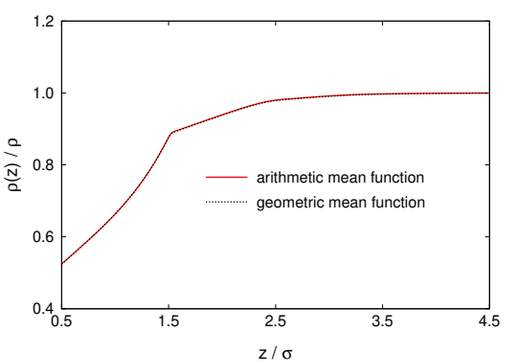

(c) Scheme 2B; $\rho^* = 0.3449$ and $\varepsilon^* = 8$




(d) Scheme 2B; $\rho^* = 0.2084$ and $\varepsilon^* = 11$




(e) Scheme 4C; $\rho^* = 0.9036$ and $\varepsilon^* = 5$




(f) Scheme 4C; $\rho^* = 0.2112$ and $\varepsilon^* = 7$

FIG. 13. Density profiles of an associating HS fluid in contact with a planar hard wall placed at $z = 0$, for the identical thermodynamic conditions presented in Section III A. The continuous red line represents our current work detailed in Section II C where an arithmetic mean function is employed to evaluate the PDF, the dash black line corresponds to the results where the arithmetic mean function has been changed with the geometric mean function that is employed in iSAFT calculation[37].