JCTC Journal of Chemical Theory and Computation

Entropy from Correlations in TIP4P Water

Emanuela Giuffré,[†] Santi Prestipino,[†] Franz Saija,[‡] A. Marco Saitta,[§] and Paolo V. Giaquinta*,[†]

Università degli Studi di Messina, Dipartimento di Fisica, Contrada Papardo, 98166 Messina, Italy, CNR − Istituto per i Processi Chimico−Fisici, Sede di Messina, Contrada Papardo, Viale Ferdinando Stagno d'Alcontres 37, 98158 Messina, Italy, and Université Pierre et Marie Curie − Paris 06, UMR 7590, IMPMC, F-75015 Paris, France

Received November 24, 2009

Abstract: We use molecular dynamics to compute the pair distribution function of liquid TIP4P water as a function of the intermolecular distance and of the five angles that are needed to specify the relative position and orientation of two water molecules. We also calculate the translational and orientational contributions to the two-body term in the multiparticle correlation expansion of the configurational entropy at three selected thermodynamic states, where we also test various approximations for the angular dependence of the pair distribution function. We finally compare the results obtained for the pair entropy of TIP4P water with the experimental values of the excess entropy of ordinary water.

I. Introduction

It is superfluous to motivate the persisting interest in a deeper understanding and more effective modeling of the equilibrium structure of water.[1] The paradigmatic status of this molecular liquid in the realms of natural and life sciences, at the interface between a variety of disciplines such as physics, chemistry, and biology, justifies the continuing efforts that are being made to refine experimental tools and theoretical approximations so as to achieve more and more reliable predictions of the microscopic and macroscopic properties of this substance as well as of the many anomalous aspects that mark its thermodynamical, structural, and dynamical behavior in a unique way.[2]

The main object of the present study is the calculation of the pair entropy of liquid water, i.e., the contribution of two-body density correlations to the configurational entropy. The statistical-mechanical framework is that provided by the multiparticle correlation expansion of the excess entropy, originally derived for a classical atomic fluid in the canonical ensemble[3,4] and later extended to the grand-canonical ensemble,[5,6] the two expressions having in fact been shown to be formally equivalent.[7] Many other authors have discussed this topic; we refer the reader to ref 8 for a short commented list of some relevant contributions to the subject. To our knowledge, the first calculation of the pair entropy of a pure molecular fluid dates back to the seminal paper authored by Lazaridis and Karplus,[9] who investigated the interplay between orientational correlations and entropy in the TIP4P model of liquid water.[10] The TIP4P potential is an effective pairwise four-site potential: a Lennard-Jones interaction site is located on the oxygen, two positive charges on the hydrogens, and an extra negative charge is located away from the oxygen along the hydrogen−hydrogen bisector. Intramolecular degrees of freedom are neglected in the TIP4P model, but nonetheless, this potential turns out to be a good transferable potential in that it ultimately provides a qualitatively correct representation of the phase diagram of water, also in comparison with other interaction models, except at very high pressure.[11,12]

The calculation of the pair entropy for a given model potential requires the knowledge of the pair distribution function (PDF) of the molecular liquid, a quantity that, in the case of TIP4P water, depends on six variables, viz., the distance between the centers of mass of the two molecules and five angles that are needed to specify their relative

* Corresponding author e-mail: paolo.giaquinta@unime.it.
[†] Università degli Studi di Messina.
[‡] CNR − Istituto per i Processi Chimico−Fisici.
[§] Université Pierre et Marie Curie.

10.1021/ct900627q
Published on Web 02/22/2010

orientation. Such a function cannot be easily obtained from numerical simulation since a statistically reliable result requires a massive numerical sampling to be carried out over a discrete six-dimensional grid whose spacing is consequently the outcome of a delicate compromise between the available computer power and the desired quantitative accuracy of the calculation. This is precisely the reason why Lazaridis and Karplus evaluated the pair entropy of liquid TIP4P water using a variety of approximations for the five-dimensional orientational distribution function (ODF) which, computed at a given intermolecular separation and then multiplied by the ordinary radial distribution function (RDF), eventually yields the full six-dimensional PDF. Such approximations implemented partial representations of the ODF on the basis of combinations of the monovariate and bivariate angular distribution functions as obtained from numerical simulation.

A few other attempts have as yet followed the route traced by Lazaridis and Karplus for the calculation of the pair entropy of water. Giaquinta and co-workers exploited one of the approximate schemes illustrated in ref 9, the so-called "adjusted gas-phase" (AGP) approximation, to highlight different ordering regimes in liquid TIP4P water[13] and to "measure" the relative amounts of positional and angular order through the translational and orientational pair entropies, used for the first time together as unbiased, self-consistent, and independent "order parameters" to map the phase diagram of water.[14] The first "exact" calculation—i.e., one performed without resorting to any approximate partial representation of the ODF—of the pair entropy was carried out by Zielkiewcz in four different models of computer water at ambient conditions.[15,16] A temperature analysis of the results for the single-point charge (SPC) model was later presented by the same author.[17] More recently, Wang and co-workers suggested a novel nonparametric approach to computing the pair entropy, alternative to the histogram-based method, and further based on a generalized Kirkwood superposition approximation (GKSA) for the ODF.[18] This approach was tested in five water models at ambient conditions.

In this paper, we present the results of a numeric calculation of the *full* PDF of liquid TIP4P water, carried out with the method of molecular dynamics (MD), at ambient conditions as well as in two other thermodynamic states, at lower temperature and higher pressure, respectively. The general theoretical framework is introduced and discussed in section II, together with the approximations that were implemented for the ODF in addition to the direct MD calculations, whose technical details are summarized in section III. The resulting translational and orientational pair entropies are presented in section IV and therein compared with the experimental data for the excess entropy of ordinary water. Section V is finally devoted to concluding remarks.

II. Theoretical Framework

A. Pair Entropy of a Molecular Liquid. Statistical mechanics provides a general expression for the entropy of a classical, atomic or molecular, fluid which, in general, can be written as an infinite sum of contributions associated with spatially integrated n-point density correlations:

$$S_{\mathrm{ex}} = \sum_{n=2}^{\infty} S_n \tag{1}$$

where S_{ex} is the excess (with respect to the corresponding ideal gas) entropy. In the absence of external fields, the two-body term—that, in the following, we shall refer to as the "pair entropy"—ordinarily delivers the dominant contribution to the excess entropy of a liquid.[13,19,20] As such, S_2 has been often used as an approximate "local" (in thermodynamic space) estimate of S_{ex} in that the calculation of the pair entropy does not require a thermodynamic potential to be integrated along an extended path connecting the state whose entropy one is interested in with another (reference) state where the thermodynamic properties of the fluid are known by independent means, as in the high-temperature ideal-gas asymptotic regime, or can be computed with other techniques.[12,21]

The pair entropy of molecular fluids reads:[9,22]

$$S_2 = -\frac{1}{2}k_{\mathrm{B}}\left(\frac{\rho}{\Omega}\right)^2 \int [g(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2) \ln g(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2) - g(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2) + 1]\, \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2\, \mathrm{d}\boldsymbol{\xi}_1\, \mathrm{d}\boldsymbol{\xi}_2 \tag{2}$$

where k_{B} is the Boltzmann constant, ρ is the particle number density, and $g(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ is the PDF which, in general, depends on the vector radii $(\mathbf{r}_1, \mathbf{r}_2)$ of the molecular centers of mass and on the pair of Euler angles sets $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$, where $\boldsymbol{\xi}_\alpha \equiv \{\theta_\alpha, \phi_\alpha, \chi_\alpha\}$ specifies the absolute orientation of the αth molecule in the laboratory reference frame $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$. The three Euler angles are respectively defined in the ranges $0 \le \theta_\alpha \le \pi$, $0 \le \phi_\alpha < 2\pi$, and $0 \le \chi_\alpha < 2\pi$, with angular elements $\mathrm{d}\boldsymbol{\xi}_\alpha = \sin(\theta_\alpha)\, \mathrm{d}\theta_\alpha\, \mathrm{d}\phi_\alpha\, \mathrm{d}\chi_\alpha$. Correspondingly, $\Omega \equiv \int \mathrm{d}\boldsymbol{\xi}_\alpha = 8\pi^2$. We note that the set of variables introduced above is in fact redundant for a homogeneous and isotropic molecular fluid since, in the absence of external fields, the PDF depends on the *relative* position and orientation of the two molecules. For water molecules, this information can be encoded into six variables only, viz., the radial separation r between the centers of mass plus five angles. In fact, let us label the oxygen and hydrogen atoms of the αth molecule as O_α, $H_{1\alpha}$, and $H_{2\alpha}$, respectively; let us further identify the mean point of the segment $H_{1\alpha}H_{2\alpha}$, joining the two hydrogen atoms, as D_α. In terms of the unit vectors

$$\hat{\mathbf{z}} = \frac{\overrightarrow{O_1O_2}}{|\overrightarrow{O_1O_2}|}, \qquad \hat{\mathbf{z}}_\alpha = \frac{\overrightarrow{O_\alpha D_\alpha}}{|\overrightarrow{O_\alpha D_\alpha}|} \tag{3}$$

one has for the angle θ_α formed by the dipole moment of the αth molecule, lying in the direction $\hat{\mathbf{z}}_\alpha$, and the intermolecular axis $\hat{\mathbf{z}}$:

$$\theta_\alpha = \arccos(\hat{\mathbf{z}}_\alpha \cdot \hat{\mathbf{z}}) \ (0 \le \theta_\alpha \le \pi) \tag{4}$$

In order to define the other Euler angles, we need to specify a reference system $\{\hat{\mathbf{x}}_\alpha, \hat{\mathbf{y}}_\alpha, \hat{\mathbf{z}}_\alpha\}$ sticking out of each molecule. This can be done by defining, for instance, the $\hat{\mathbf{x}}_\alpha$ and $\hat{\mathbf{y}}_\alpha$ axes as

$$\hat{\mathbf{x}}_\alpha = \frac{\overrightarrow{H_{2\alpha}H_{1\alpha}}}{|\overrightarrow{H_{2\alpha}H_{1\alpha}}|}, \qquad \hat{\mathbf{y}}_\alpha = \hat{\mathbf{z}}_\alpha \wedge \hat{\mathbf{x}}_\alpha \tag{5}$$

Hence, in terms of the auxiliary unit vector

$$\hat{\mathbf{N}}_\alpha \equiv \frac{\hat{\mathbf{z}} \wedge \hat{\mathbf{z}}_\alpha}{|\hat{\mathbf{z}} \wedge \hat{\mathbf{z}}_\alpha|} \tag{6}$$

specifying the line of nodes that is associated with the αth molecule, we obtain

$$\phi_\alpha = \arg(\hat{\mathbf{x}} \cdot \hat{\mathbf{N}}_\alpha, \hat{\mathbf{y}} \cdot \hat{\mathbf{N}}_\alpha) \ (0 \le \phi_\alpha < 2\pi) \tag{7}$$

and

$$\chi_\alpha = \arg(\hat{\mathbf{x}}_\alpha \cdot \hat{\mathbf{N}}_\alpha, -\hat{\mathbf{y}}_\alpha \cdot \hat{\mathbf{N}}_\alpha) \ (0 \le \chi_\alpha < 2\pi) \tag{8}$$

where $\arg(u, v)$ is the polar argument of the vector (u, v). The angle χ_α describes the rotation of the αth water molecule around its own dipole vector. One can readily see that one of the above six angles is superfluous; in fact, upon choosing, say, $\hat{\mathbf{x}} \equiv \hat{\mathbf{N}}_1$, the angle ϕ_1 identically vanishes. In the following, we shall refer to this set of angular coordinates as the *A* set. Actually, one may further take advantage of a symmetry of the PDF that follows from the observation that the pointing directions of the unit vectors $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ depend on the (arbitrary) way one has labeled the two hydrogen atoms within each water molecule. In turn, this choice uniquely determines the pointing directions of $\hat{\mathbf{y}}_1$ and $\hat{\mathbf{y}}_2$. This residual freedom reflects on the behavior of the (exact) PDF that is in fact invariant under χ_α rotations moving this angle to $[\chi_\alpha + \pi(\mathrm{mod}\ 2\pi)]$. Hence, it should suffice to consider the dependence of the PDF on χ_α within the interval $[0, \pi]$. We shall refer to this modified—by effect of symmetry—set of variables as the A_s set.

The above choice of variables closely recalls the one made by Lazaridis and Karplus.[9] However, in addition to the pairs (θ_1, θ_2) and (χ_1, χ_2), they used the angle $\phi \equiv \phi_1 - \phi_2$ describing the relative rotation of a pair of molecules (more specifically, of their respective line-of-nodes unit vectors, $\hat{\mathbf{N}}_1$ and $\hat{\mathbf{N}}_2$) around the intermolecular axis. Moreover, their conventions on how to measure the two pairs of angles mentioned above differ from those specified for the set *A*. In particular, H_{11} was explicitly taken to be the hydrogen atom of molecule 1 that is closest to the oxygen atom of molecule 2, and similarly for H_{21} (see Figure 1 of ref 9 and the relative caption). Lazaridis and Karplus further exploited a number of symmetries of the PDF involving all five angles. In fact, in addition to those introduced above which concern the pair (χ_1, χ_2), they also considered two residual symmetries of the PDF: they noted that the two water molecules are interchangeable, which allows one to integrate the angle θ_2 from θ_1 to π, and also observed that the angle ϕ can be integrated from 0 to π only. We shall refer to this set of angular coordinates as the B_s set which, however, we shall implement without resorting to the additional symmetry concerning the angle θ_2.

A still different choice (*C*) was made by Zielkiewicz,[15,16] who assumed $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\} \equiv \{\hat{\mathbf{x}}_1, \hat{\mathbf{y}}_1, \hat{\mathbf{z}}_1\}$, the orthonormal triad attached to molecule 1 being specified as above by eqs 3 and 5. He then considered the spherical coordinates of O_2 in the reference frame centered on O_1:

$$\Theta = \arccos(\hat{\mathbf{r}} \cdot \hat{\mathbf{z}}_1), \qquad \Phi = \arg(\hat{\mathbf{x}}_1 \cdot \hat{\mathbf{r}}_\perp, \hat{\mathbf{y}}_1 \cdot \hat{\mathbf{r}}_\perp) \tag{9}$$

where $\mathbf{r} \equiv \overrightarrow{O_1O_2}$ and $\mathbf{r}_\perp = \mathbf{r} - (\mathbf{r} \cdot \hat{\mathbf{z}}_1)\,\hat{\mathbf{z}}_1$, with $0 \le \Theta \le \pi$ and $0 \le \Phi < 2\pi$, in addition to the three Euler angles $(\theta_2, \phi_2, \chi_2)$ that specify the global orientation of the second molecule with respect to the central one, according to eqs 4, 7, and 8. As observed before, the alternative for $\hat{\mathbf{x}}_\alpha$ is 2-fold, allowing one to restrict the range of ϕ_2 and χ_2 to the interval $[0, \pi]$. The set of variables where this latter symmetry property is employed will be referred to as the C_s set.

It is important to realize that different choices of the angles that specify the relative orientation of two water molecules, as well as of the associated intervals within which the PDF is then being sampled, lead to different histograms and, consequently, to potentially different numerical estimates of the pair entropy. This is obviously due to the finite integration meshes as well as to the number of symmetries that are implemented in each angular set. In fact, we should expect that the best results will be achieved when one makes full use of the symmetries of the PDF, which is actually the case of the B_s set. In this respect, we note that, for an assigned number of configurations over which the PDF histogram is being sampled, the information contents of the histograms built by using the A_s and C_s sets should be four times larger than that of the corresponding histograms in the *A* and *C* sets. On the other hand, the statistical quality achieved with the B_s set should be twice as good as that achieved with the A_s and C_s sets.

Now, let $\mathbf{\Psi}$ be any 5-fold set of independent angular variables. Upon observing that $\int \mathrm{d}\boldsymbol{\xi}_1\, \mathrm{d}\boldsymbol{\xi}_2 = 2\pi \int \mathrm{d}\mathbf{\Psi}$, we can rewrite eq 2 as

$$s_2 = -k_\mathrm{B}\left(\frac{2\pi}{\Omega}\right)^2 \rho \int [g(r, \mathbf{\Psi}) \ln g(r, \mathbf{\Psi}) - g(r, \mathbf{\Psi}) + 1] r^2\, \mathrm{d}r\, \mathrm{d}\mathbf{\Psi} \tag{10}$$

where s_2 is the pair entropy per particle. Following Lazaridis and Karplus,[9] we factorize the PDF as

$$g(r, \mathbf{\Psi}) = g(r) g(\mathbf{\Psi}|r) \tag{11}$$

where $g(r)$ is the RDF and $g(\mathbf{\Psi}|r)$ is the conditional distribution for the relative orientation of two molecules separated by a distance *r*, i.e., the quantity that we shall refer to in the following as the ODF. The following normalization condition holds:

$$\frac{2\pi}{\Omega^2} \int g(r, \mathbf{\Psi})\, \mathrm{d}\mathbf{\Psi} = g(r) \tag{12}$$

Correspondingly, the pair entropy can be resolved into the sum of two terms:

$$s_2 = s_2^{(\mathrm{tr})} + s_2^{(\mathrm{or})} \tag{13}$$

where

$$s_2^{(\mathrm{tr})} = -k_\mathrm{B}(2\pi)\rho \int_0^\infty [g(r) \ln g(r) - g(r) + 1] r^2\, \mathrm{d}r \tag{14}$$

is the translational pair entropy and

$$s_2^{(or)} = \int_0^\infty \rho\, g(r) S^{(or)}(r) r^2 dr \quad (15)$$

is the orientational pair entropy that is defined in terms of the orientational local entropy (OLE):

$$S^{(or)}(r) = -k_B \left(\frac{2\pi}{\Omega}\right)^2 \int g(\mathbf{\Psi}|r) \ln g(\mathbf{\Psi}|r)\, d\mathbf{\Psi} \quad (16)$$

B. Approximations for the Pair Distribution Function. At variance with the RDF, the numerical evaluation of the ODF is a formidable computational task because of the number of variables this function depends on. Approximating the ODF may be in many cases the only viable route to the calculation of the orientational contribution to the thermodynamic properties of liquid water. In this respect, Lazaridis and Karplus proposed a number of approximate schemes for the full PDF, essentially based on the assumption that the ODFs of gaseous and liquid water have similar short-range structures.[9] The first "family" of such approximations was generated by factorizing the ODF into a product of one-dimensional (1d) and two-dimensional (2d) marginals, defined as the probability distributions of one angle, or the joint probability distributions of two angles, regardless of the values attained by the remaining angles:

$$g(\Psi_i|r) \equiv \frac{\int g(\mathbf{\Psi}|r)\, d^4\Psi_{j\neq i}}{\int d^4\Psi_{j\neq i}} \quad (17)$$

and

$$g(\Psi_i, \Psi_j|r) \equiv \frac{\int g(\mathbf{\Psi}|r)\, d^3\Psi_{k\neq i,j}}{\int d^3\Psi_{k\neq i,j}} \quad (18)$$

Lazaridis and Karplus (LK) introduced and discussed various factorizations that were tested against the thermodynamic properties (energy, entropy) of the gas phase of TIP4P water.[9] They eventually concluded that, overall, the most balanced scheme was the one which they referred to as the F7 factorization:

$$g_{LK}^{(F7)}(\theta_1, \theta_2, \phi, \chi_1, \chi_2) \equiv \left[\frac{g(\theta_1, \theta_2)\, g(\chi_1, \chi_2)\, g(\theta_1, \chi_2)\, g(\theta_2, \chi_1)}{g(\theta_1)\, g(\theta_2)\, g(\chi_1)\, g(\chi_2)}\right] g(\phi) \quad (19)$$

where the parametric dependence on the radial separation r is implicit in both the full and marginal ODFs that appear on the left- and right-hand sides of eq 19, respectively. This approximation exploits the "flatness" of the ODF with respect to the angle ϕ and the ensuing absence, in the liquid phase, of significant correlations between ϕ and the other angles.[9]

In addition to the factorization scheme illustrated above, Lazaridis and Karplus described a different approach based on the low-density limit of the ODF, modified by a subset of the marginal ODFs that were obtained from the numerical simulations of *liquid* TIP4P water. The resulting AGP approximation for the ODF can be written as follows:

$$g_{LK}^{(AGP)}(\theta_1, \theta_2, \phi, \chi_1, \chi_2) = g_s(\theta_1, \theta_2, \phi, \chi_1, \chi_2) \cdot C_{LK}(\theta_1, \theta_2, \phi, \chi_1, \chi_2) \quad (20)$$

where

$$C_{LK}(\theta_1, \theta_2, \phi, \chi_1, \chi_2) = \frac{g(\theta_1, \theta_2)}{g_s(\theta_1, \theta_2)} \cdot \frac{g(\chi_1, \chi_2)}{g_s(\chi_1, \chi_2)} \cdot \frac{g(\phi)}{g_s(\phi)} \quad (21)$$

and where we have again suppressed the dependence on the intermolecular distance r. In eq 21, $g(\theta_1, \theta_2)$, $g(\chi_1, \chi_2)$, and $g(\phi)$ are the "exact" marginals, which can be calculated by numerical simulation, while the function g_s and the associated marginals follow from

$$g_s(\mathbf{\Psi}|r) \equiv g_{gas}(\mathbf{\Psi}|r) + \mathcal{J}(r)[1 - g_{gas}(\mathbf{\Psi}|r)] \quad (22)$$

where

$$g_{gas}(\mathbf{\Psi}|r) = \left(\frac{\Omega^2}{2\pi}\right) \frac{\exp[-\beta u(r, \mathbf{\Psi})]}{\int \exp[-\beta u(r, \mathbf{\Psi})]\, d\mathbf{\Psi}} \quad (23)$$

with $u(r, \mathbf{\Psi})$ being the interaction potential and

$$\mathcal{J}(r) = \frac{U_{MD}(r) - U_{gas}(r)}{U(r) - U_{gas}(r)} \quad (24)$$

In eq 24, $U_{MD}(r)$ is the angularly averaged interaction energy evaluated at a specified distance r through the MD simulation, $U_{gas}(r)$ is the corresponding quantity calculated in the gas phase, and $U(r)$ is the unweighted average of $u(r, \mathbf{\Psi})$ over orientations. The ODF that follows from eq 20 must be eventually normalized so as to satisfy eq 12. The smoothing function $\mathcal{J}(r)$ is such that the angularly averaged energy evaluated through the function g_s reproduces the quantity $U_{MD}(r)$.

As seen from eqs 20 and 21, the naive gas-phase approximation—formulated in eqs 22, 23, and 24—is *adjusted* so as to enforce, through the C_{LK} factor, a reasonable degree of consistency between the resulting ODF and a subset of liquid-phase marginal distributions. Wang and co-workers[18] recently noted that a non-negligible correlation exists between the angles θ and χ of each water molecule which cannot be accounted for by the mere product of the corresponding 1d marginals, i.e., $g(\theta)$ and $g(\chi)$. Hence, they proposed a GKSA factorization of the ODF that differs from the F7 scheme set up by Lazaridis and Karplus in that the modified ODF includes an extra pair of 2d "intramolecular" marginals, viz., $g(\theta_1, \chi_1)$ and $g(\theta_2, \chi_2)$. Motivated by this observation, we also report in this paper on three variants of the original AGP approximation where the coupling factor C_{LK}, which appears in eq 20, has been replaced, in turn, by the following expressions:

$$C_1(\theta_1, \theta_2, \phi, \chi_1, \chi_2) = C_{LK}(\theta_1, \theta_2, \phi, \chi_1, \chi_2) \left[\frac{g(\theta_1, \chi_1)}{g_s(\theta_1, \chi_1)} \cdot \frac{g(\theta_2, \chi_2)}{g_s(\theta_2, \chi_2)}\right] \quad (25)$$

$$C_2(\theta_1,\theta_2,\phi,\chi_1,\chi_2) = C_{LK}(\theta_1,\theta_2,\phi,\chi_1,\chi_2) \times \left[\frac{g(\theta_1,\chi_2)}{g_s(\theta_1,\chi_2)} \cdot \frac{g(\theta_2,\chi_1)}{g_s(\theta_2,\chi_1)}\right] \quad (26)$$

$$C_3(\theta_1,\theta_2,\phi,\chi_1,\chi_2) = C_2(\theta_1,\theta_2,\phi,\chi_1,\chi_2) \times \left[\frac{g(\theta_1,\chi_1)}{g_s(\theta_1,\chi_1)} \cdot \frac{g(\theta_2,\chi_2)}{g_s(\theta_2,\chi_2)}\right] \quad (27)$$

We shall refer to the three modified AGP schemes reported in eqs 25, 26, and 27 as the AGP1, AGP2, and AGP3 approximations, respectively. We observe that the AGP2 approximation includes the same 2d marginals that also appear in the F7 approximation originally proposed by Lazaridis and Karplus,[9] while the AGP3 coupling factor reproduces the factorization of 2d marginals adopted by Wang and co-workers.[18]

In order to simplify the calculation of the pair entropy, Lazaridis and Karplus further resorted to an approximate computational strategy. In fact, they averaged the liquid-phase angular distributions that appear in both the F7 and AGP schemes over three distinct spatial regions ("shells") whose boundaries were chosen so as to coincide with the positions of the first peak and of the first two troughs in the oxygen−oxygen RDF of TIP4P water at 25 °C and 1 atm. Their choice leads to the following intervals: $0 \le r \le 0.28$ nm, $0.28 \text{ nm} < r \le 0.34$ nm, and $0.34 \text{ nm} < r \le 0.56$ nm. As for the three gas-phase marginals that also contribute to the AGP approximation, they were apparently calculated at three "representative" distances corresponding to the three regions over which the same marginals were calculated in the liquid by simulation, i.e, 0.28 nm, 0.32 nm, and 0.45 nm. We note that a similar computational strategy was adopted by Wang and co-workers[18] who also computed the OLE over three consecutive shells that closely correspond to the analogous choice made by Lazaridis and Karplus. However, they used a different approach to estimate the pair entropy, viz., the so-called *k*th nearest-neighbor method.

The AGP approximation clearly depends on the preknowledge of the RDF spatial profile of the model in a given thermodynamic state. We propose a generalization of this method by averaging the liquid-phase marginals over a variable number of shells, $N_{shell} = R_{max}/\Delta R_{shell}$, where R_{max} is the chosen distance cutoff and ΔR_{shell} is the width of each shell. Obviously, the maximum number of shells over which the marginals can be computed corresponds to a discretization of the interval such that $\Delta R_{shell} = \Delta r$, where Δr is the spatial resolution of the calculation. In addition, the gas-phase marginals are computed at the midpoint of each interval. We shall refer to this modified implementation of the AGP approximation as the multishell AGP (MSAGP) approximation.

III. Simulation Method and Technical Details

We carried out constant-pressure constant-temperature MD simulations of the TIP4P model, implemented through the PINY code.[23] The system contained 512 molecules in a cubic cell with periodic boundary conditions. The spherical cutoff of the Lennard-Jones interactions was 0.8 nm. Electrostatic interactions were modeled with the particle-mesh Ewald method. A time step of 2.5 fs turned out to be sufficient for a proper dynamical evolution since the TIP4P model is based on a rigid-molecule description. Typical simulation times were in the 50 ns range, corresponding to runs 2×10^7 steps long. The MD configurations were stored every 0.5 ps in such a way that the PDF was calculated over as many as 10^5 system snapshots. Under ambient conditions, the molecular density was initially 1 g cm^{-3}, corresponding to a width of the simulation cell of 2.48 nm.

As already emphasized, the evaluation of the pair entropy from eq 10 is far from being a trivial task, it being the outcome of a six-dimensional integration. In practice, many aspects of the calculation may seriously influence the final numerical accuracy such as (i) the number, N_{conf}, of configurations contributing to the thermal averages, a parameter that affects the "quality" of the integrand, i.e., its regularity as a function of the independent variables; (ii) the integration mesh sizes (Δr, $\Delta\mathbf{\Psi}$); (iii) the numerical integration method, whose impact on the result is not *a priori* obvious, especially when the integrand happens to be noisy.

The RDF and the full PDF were respectively estimated through the following formulas:

$$g(r) = I(r;\Delta r)\left\{\frac{4\pi}{3}\rho\left[\left(r+\frac{\Delta r}{2}\right)^3 - \left(r-\frac{\Delta r}{2}\right)^3\right]\right\}^{-1} \quad (28)$$

$$g(r,\mathbf{\Psi}) = I(r,\mathbf{\Psi};\Delta r,\Delta\mathbf{\Psi})\left\{\frac{4\pi}{3}\rho\left[\left(r+\frac{\Delta r}{2}\right)^3 - \left(r-\frac{\Delta r}{2}\right)^3\right]\right\}^{-1}\left(\frac{2\pi}{\Omega^2}\Delta\mathbf{\Psi}\right)^{-1} \quad (29)$$

where $I(r;\Delta r)$ and $I(r,\mathbf{\Psi};\Delta r,\Delta\mathbf{\Psi})$ are the corresponding histograms and

$$\Delta\mathbf{\Psi} = \left[-\cos\left(\theta_1+\frac{\Delta\theta_1}{2}\right)+\cos\left(\theta_1-\frac{\Delta\theta_1}{2}\right)\right]\times \left[-\cos\left(\theta_2+\frac{\Delta\theta_2}{2}\right)+\cos\left(\theta_2-\frac{\Delta\theta_2}{2}\right)\right]\times \Delta\chi_1\Delta\chi_2\Delta\phi_2 \quad (30)$$

We also took advantage of the symmetry properties of $I(r,\mathbf{\Psi};\Delta r,\Delta\mathbf{\Psi})$, with the effect of doubling the statistics whenever the range of a given angle is halved from 2π to π. We assumed $\Delta r = 0.01$ nm and $\Delta\Psi_i = 10°$ for each of the five angles that were involved in the calculation. This latter value turned out to be a reasonable compromise between histogram resolution and statistics, and moreover, it is the value that has been commonly used in the past literature on the subject.[9,15,18] All the integrations were performed using the standard Simpson method. In order to get a quantitative feeling on the numerical error associated with the choices made for Δr and $\Delta\Psi_i$, we compared the RDF obtained directly from the simulation with the output from the normalization condition reported in eq 12 in all the thermodynamic states investigated. We found that the relative deviation was about 3% at very short distances, in the region corresponding to the initial rise of the RDF from zero up to its highest maximum, and rapidly dropped to zero

Table 1. Translational and Orientational Pair Entropies (e.u.) of TIP4P Water at 298 K[a]

source	$s_2^{(tr)}$	$s_2^{(or)}$	s_2
Lazaridis and Karplus[9]	−3.14	−9.1[b]	−12.2
Lazaridis and Karplus[9]	−3.14	−11.7[c]	−14.8
Wang et al.[18]	−3.15	−10.52	−13.67
Zielkiewicz[15,16]	−2.97[d]	−11.9[e]	−14.9[f]

[a] Data from refs 9 and 18 refer to constant-pressure simulations carried out at 1 atm, while data from refs 15 and 16 refer to constant-volume simulations corresponding to a density ρ_m = 0.999 g cm^{-3}. [b] Estimate obtained using the F7 approximation. [c] Estimate obtained using the AGP approximation. [d] The estimate originally provided in ref 15 has been corrected by adding the missing contribution reported in ref 16. [e] Estimate inferred upon subtracting the translational pair entropy from the cumulative pair entropy. [f] Estimate inferred upon subtracting the ideal-gas entropy from the absolute entropy reported in ref 16.

with increasing distances, already being less than 0.1% for $r \geq 0.3$ nm.

IV. Results

We investigated the properties of TIP4P water in three thermodynamic states located in the stable liquid-phase region, i.e., (T = 260 K, P = 1 bar), (T = 300 K, P = 1 bar), and (T = 300 K, P = 4 kbar). We shall first present the results obtained at ambient conditions, which we shall compare with those obtained by other authors using the Monte Carlo or the molecular dynamics method, either resorting to approximate representations of the ODF[9,18] or carrying out a full direct calculation.[15,16] We shall then illustrate how the pair entropy changes upon lowering the temperature down to 260 K or increasing the pressure up to 4 kbar.

A. TIP4P Water at Ambient Conditions. We computed equilibrium averages at ambient conditions (T = 300 K, P = 1 bar) over sets of up to 10^5 MD configurations. The resulting average values of the specific density ρ_m and of the excess internal energy U_{ex} were 0.986 g cm^{-3} and −10.01 kcal mol^{-1}, respectively. Under such thermodynamic conditions, the cumulative ideal-gas contribution to the entropy of TIP4P water, modeled as a gas of noninteracting rigid molecules, amounts to 30.74 entropy units (e.u.; 1 e.u. = 4.184 J K^{-1} mol^{-1}),[24] about one-third of which (10.44 e.u.) is to be ascribed to the rotational degrees of freedom. We report in Table 1 the currently available estimates of the translational, orientational, and cumulative pair entropies of TIP4P water at ambient conditions for a comparison with the present results.

The translational pair entropy can be computed in a straightforward way since it requires the RDF only as an input (see Figure 1). Upon using eq 14, we obtained −3.05 e.u. for this quantity.

1. MD Results for the Orientational Pair Entropy. The OLE computed directly from simulation—i.e., using no approximate partial representation—is shown in Figure 2 for all the angular sets introduced in section II.B. The discrepancies observed between some of the five estimates are almost entirely due to their different statistical qualities, as already discussed in section IIB. This aspect is clarified in Figure 3 where we reported the OLEs obtained using the

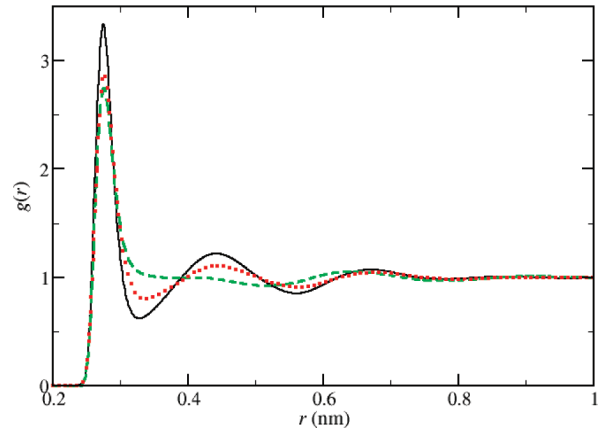

Figure 1. Radial distribution function of TIP4P water averaged over 10^5 configurations: (T = 260 K, P = 1 bar), black continuous curve; (T = 300 K, P = 1 bar), red dotted curve; (T = 300 K, P = 4 kbar), green dashed curve.

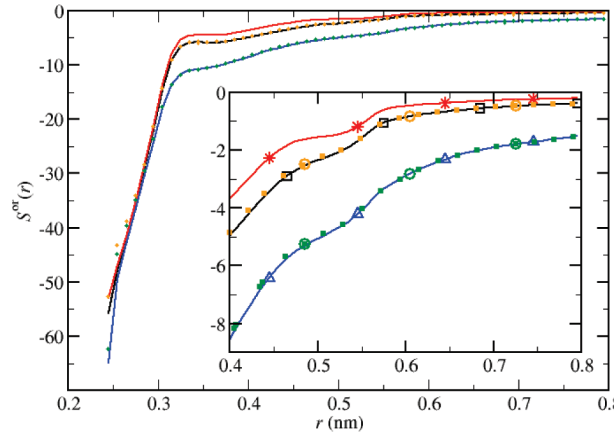

Figure 2. Orientational local entropy (e.u.) of TIP4P water at ambient conditions averaged over 10^5 MD configurations: *A* set, blue solid curve (further marked with triangles in the inset); *C* set, green dotted curve (further marked with circles in the inset); A_s set, black solid curve (further marked with squares in the inset); C_s set, orange dotted curve (further marked with circles in the inset); B_s set, red solid curve (further marked with asterisks in the inset).

(A, C), (A_s, C_s), and B_s angular sets, the corresponding histograms now being averaged over sets of configurations whose numbers lie in the ratio 8:2:1, respectively, as suggested by the number of symmetries employed in each angular set. As a result, the five estimates are manifestly seen to collapse onto the same curve. This comparison confirms that the B_s set provides the most accurate estimate that can be generated with an assigned number of configurations.

Figure 4 shows the function $I_S^{(or)}(r) = \rho\, g(r)\, S^{(or)}(r)\, r^2$, that yields upon integration the orientational pair entropy (see eq 15), depicted for increasing values of N_{conf}. The most relevant feature of this function is the profound minimum located at r = 2.75 nm, i.e., where the RDF attains its maximum value corresponding to the first coordination shell. We observe that the shape of the minimum, aside from its depth, does not appreciably change with the number of configurations; on the other hand, the longer-range part of the function shifts almost rigidly toward zero as N_{conf} increases. We found that, even upon sampling the $I_S^{(or)}(r)$

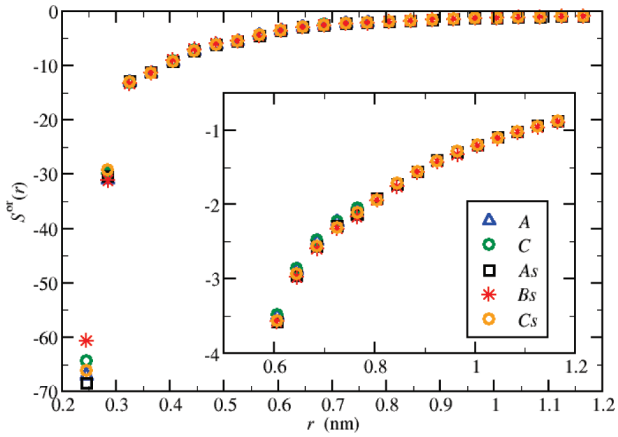

Figure 3. Orientational local entropy (e.u.) of TIP4P water at ambient conditions averaged over 8×10^4 configurations for the *A* and *C* sets, over 2×10^4 configurations for the A_s and C_s sets, and over 1×10^4 configurations for the B_s set.

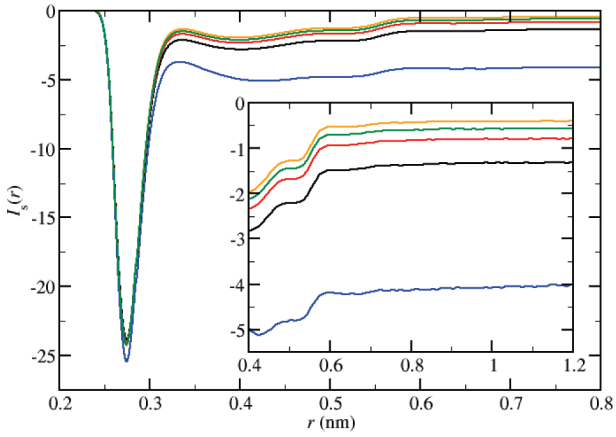

Figure 4. The function $I_S^{(or)}(r) = \rho\, g(r)\, S^{(or)}(r)\, r^2$, computed for the B_s angular set at ambient conditions, plotted as a function of the distance for increasing values of the number of configurations $N_{conf} = \{1, 3, 5, 7, 10\} \times 10^4$. The inset shows the long-range behavior of the function, decreasing in absolute terms with increasing N_{conf} at fixed *r*.

histogram over 10^5 configurations, the function has not decayed yet to zero over distances corresponding to half the width of the simulation cell. This behavior suggests that, over such intermolecular separations, the decorrelation time of the angular degrees of freedom is on the order of the time (~50 ns) spanned by the longest MD trajectory generated in the present calculations. The concurring effect of the finite mesh size cannot be excluded as well. Upon integrating the most refined histogram produced for $I_S^{(or)}(r)$, we obtained −14.7 e.u. for the orientational pair entropy, a value that is certainly underestimated because of the arguments illustrated before. Hence, in order to obtain a more reliable estimate, we resorted to an extrapolation of the quantity $\sigma(r; N_{conf}) \equiv g(r; N_{conf})\, S^{(or)}(r; N_{conf})$, that was modeled, as a function of N_{conf}, according to the following inverse-power law:

$$\sigma(r;N_{conf}) = \tilde{\sigma}(r) + \frac{A(r)}{N_{conf}^{\alpha(r)}} \qquad (31)$$

where $\tilde{\sigma}(r)$, $A(r)$, and $\alpha(r)$ are *r*-dependent parameters that were determined through a least-squares fit of the MD data.

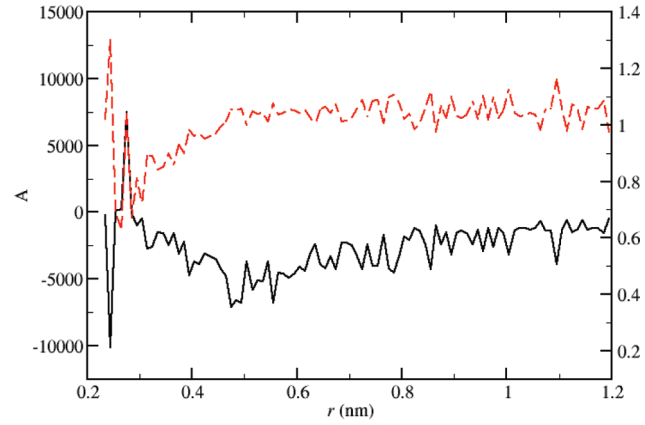

Figure 5. Space dependent amplitude (continuous black curve, left axis) and exponent (broken red curve, right axis) of the extrapolating function used for $\sigma(r, N_{conf})$ in the B_s set at ambient conditions.

This procedure differs in a significant way from the one used in ref 15, where, instead, a series of estimates of the absolute entropy, obtained over MD trajectories of increasing length, was extrapolated as a function of the simulation time. However, as noted above, such estimates are likely affected by a truncation error arising from the nonvanishing tail of the OLE. This is the reason why we extrapolated this function, and only afterward performed the integration in eq 15.

The current fit was carried out over seven sets of data corresponding to values of N_{conf} ranging between 4×10^4 and 10×10^4, with sequential increments of 10^4 configurations. Moreover, every set was assigned a weight proportional to the number of configurations used to calculate the function $\sigma(r; N_{conf})$. The resulting best-fit values of the parameters $A(r)$ and $\alpha(r)$ are plotted in Figure 5 as a function of *r*. Notwithstanding their apparently noisy aspect, the quality of the fit was good everywhere: in fact, the minimum root-mean-square deviation from the MD data turned out to be less than 10^{-5} e.u. in the region of the first minimum—where, as noted before, the OLE does not exhibit a marked dependence on N_{conf} —decreasing further with distance by more than 2 orders of magnitude for $r > 0.6$ nm. The long-range trend of $\alpha(r)$ indicates that the tail of $\sigma(r; N_{conf})$ deviates from $\tilde{\sigma}(r)$ approximately as N_{conf}^{-1}. The asymptotic estimate of the integrand function, $\tilde{I}_S^{(or)}(r)$, that we obtained for the B_s set is depicted in Figure 6. Upon comparing this result with the present largest N_{conf} estimate of the same quantity, one notices that the first deep minimum was not significantly affected by the extrapolation, whereas the medium- and long-range behavior actually was to an appreciable extent. In fact, at variance with the $N_{conf} = 10^5$ estimate which flattens off over the largest sampled distances at a value of about −0.4 entropy units, the extrapolated function has already decayed to zero over distances beyond the third coordination shell. As for the resulting orientational pair entropy, we obtained—upon integrating $\tilde{I}_S^{(or)}(r)$—the value −11.5 e.u., an estimate that is fairly close to the one (−11.9 e.u.) that we inferred from the results reported in refs 15 and 16 at $T = 298$ K, which also were obtained without resorting to any approximate partial representation of the ODF. In any case, a

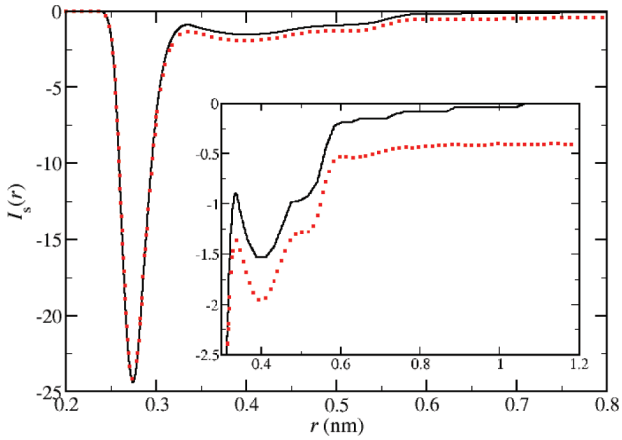

Figure 6. Integrand function, $I_S^{(gr)}(r)$, computed in the B_s set at ambient conditions: continuous (black) curve, asymptotic ($N_{conf} \rightarrow \infty$) estimate; dotted (red) curve, $N_{conf} = 10^5$ estimate.

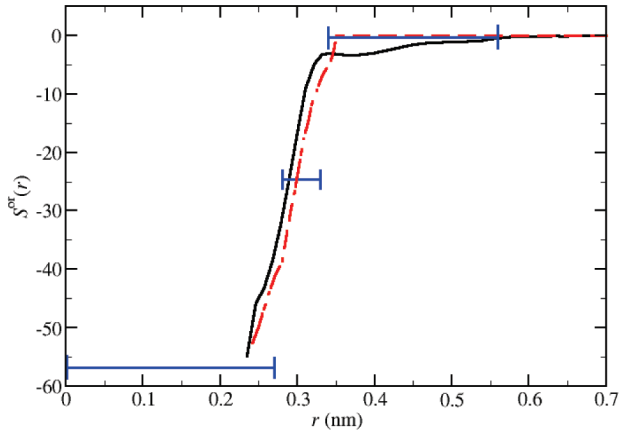

Figure 7. Orientational local entropy at ambient conditions: continuous (black) curve, present asymptotic estimate obtained in the B_s set; dot-dashed (red) curve, approximate three-shell AGP estimate (reproduced from Figure 4 of ref 9); horizontal (blue) bars, approximate three-shell GKSA estimates from ref 18.

modest increase of the orientational pair entropy such as the one we registered (∼0.4 e.u.) might be consistent with the 2 K temperature gap between the two calculations.

2. Approximate Results for the Orientational Pair Entropy. Figure 7 shows the outcome of our asymptotic extrapolation of $S^{(or)}(r)$ that is compared with the corresponding approximate estimate obtained by Lazaridis and Karplus using the AGP approximation.[9] We also included the OLE values (−56.77, −24.58, −0.34) that were obtained, in entropy units, from the orientational Shannon entropies of TIP4P water as calculated by Wang and co-workers (see Table 1 of ref 18) over three "representative" shells, i.e., $0 \leq r \leq 0.27$, $0.27 < r \leq 0.33$, and $0.33 < r \leq 0.56$, all distances being expressed in nanometers. We observe that both approximations miss the weakly modulated medium range tail of $S^{(or)}(r)$. Indeed, the modest global shift toward larger distances observed in the Lazaridis and Karplus estimate, as compared with the present calculation, clearly compensates for the faster and abrupt decay to zero of the function which, upon integration, does in fact lead to a result

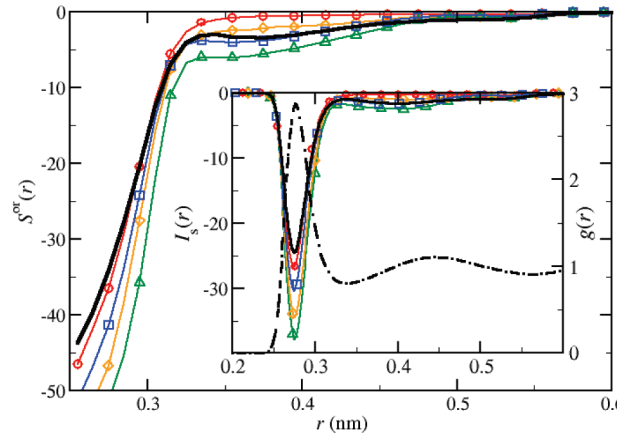

Figure 8. Orientational local entropy at ambient conditions: black curve, present asymptotic estimate obtained in the B_s set; red curve marked with circles, MSAGP estimate; blue curve marked with squares, MSAGP1 estimate; orange curve marked with diamonds, MSAGP2 estimate; green curve marked with triangles, MSAGP3 estimate. The inset shows the corresponding integrand functions $I_S^{(or)}(r)$ and the radial distribution function sampled over 10^5 configurations (right axis).

for the orientational pair entropy that is very close to the current estimate (see Table 1). A similar performance of the three-shell AGP approximation, with a partial yet fortuitous error compensation, was also observed in the other two thermodynamic states that we shall discuss in the following sections. On the other hand, the estimate reported by Wang and co-workers is one entropy unit larger than the present one, a discrepancy that is less obvious to explain since the method used by these latter authors was not based on a partial or full calculation of the ODF histogram, as done in ref 9 as well as in the present work.

The OLEs generated by the highest resolution ($\Delta R_{max} = \Delta r$) multishell variants of the AGP approximation introduced in section II.B are shown in Figure 8 for a reduced set of configurations ($N_{conf} = 10^4$). We first note that the multishell implementation of the original AGP approximation definitely improves over the original three-shell AGP approximation at short distances. In fact, this more spatially refined estimate closely reproduces the rise of the function for increasing r: the curve neither shows the rigid shift that we commented on above nor the kink at $r = 0.34$ nm, which is actually an artifact of using marginals averaged over discrete regions. Note, however, that the MSAGP approximation still fails to reproduce the tail of the function. Somewhat paradoxically, notwithstanding the improvement observed in the OLE at short distances, the multishell implementation of the AGP approximation leads to a lower estimate (in absolute value) of the orientational pair entropy (see Table 2) as compared with that obtained from the corresponding three-shell version (see Table 1). The worsened agreement is the consequence of the no longer fortuitously compensating failures which affected the short-range and long-range parts of $S^{(or)}(r)$ as estimated by Lazaridis and Karplus.

Better results on the long-range decay can be obtained by suitably modifying the AGP approximation so as to take into account missing angular correlations that plausibly play a

Table 2. Multishell AGP Estimates of the Orientational Pair Entropy (e.u.) of TIP4P Water Obtained, at Different Temperatures and Pressures, upon Sampling 1d and 2d Marginals over 10^4 Configurations, Using the Largest Number of Shells Compatible with the Spatial Resolution of the Calculation ($\Delta R_{max} = \Delta r = 0.01$ nm)[a]

approximation	I[b]	II[c]	III[d]
MSAGP	−11.7	−9.6	−9.0
MSAGP1	−15.8	−12.6	−12.7
MSAGP2	−16.6	−13.6	−13.0
MSAGP3	−21.2	−17.0	−17.4
Simulation	−14.8	−11.5	−11.3

[a] The asymptotic estimates generated by the extrapolated simulation data are also included for comparison. [b] $T = 260$ K, $P = 1$ bar ($R_{max} = 1.20$ nm). [c] $T = 300$ K, $P = 1$ bar ($R_{max} = 1.20$ nm). [d] $T = 300$ K, $P = 4$ kbar ($R_{max} = 1.15$ nm).

Table 3. Convergence of the Multishell AGP Approximations with Increasing Numbers of Configurations at $T = 300$ K and $P = 1$ bar

	number of configurations		
approximation	1×10^3	5×10^3	10×10^3
MSAGP	−9.74	−9.64	−9.62
MSAGP1	−12.91	−12.69	−12.63
MSAGP2	−13.82	−13.63	−13.60
MSAGP3	−17.44	−17.09	−17.02

role at larger distances. As discussed in section II.B, we tested three differently augmented AGP schemes. The results are shown in Figure 8. It clearly emerges that the MSAGP1 approximation is closer to the asymptotic MD estimate, from which it significantly deviates at very short distances only. The deeper minimum in $I_S^{(or)}(r)$ substantially accounts for the 10% discrepancy between the MSAGP1 and simulation estimates of $s_2^{(or)}$ (see Table 2). We recall that this particular scheme includes the *intramolecular* correlation between the angles θ and χ through the marginal distributions $g(\theta_1, \chi_1)$ and $g(\theta_2, \chi_2)$ whose role and importance have been recently highlighted by Wang and co-workers.[18] However, their approximation as well as the F7 factorization scheme also include cross-*intermolecular* correlations between the same angular pairs, a combination that we also exploited in the MSAGP3 approximation. As seen from Figure 8, the inclusion of the marginal distributions $g(\theta_1, \chi_2)$ and $g(\theta_2, \chi_1)$ worsens the agreement with the simulation data in that these additional correlations do actually overemphasize the structure of $S^{(or)}(r)$ both at small and large distances, producing an even more profound minimum in $I_S^{(or)}(r)$ as well as a more prominent long-range tail. Correspondingly, the estimated orientational pair entropy drops by more than four entropy units. The major responsibility of intermolecular (θ, χ) correlations in hollowing out a deeper minimum in the integrand function is also confirmed by the outcome of the MSAGP2 approximation where intramolecular (θ, χ) correlations have been neglected.

A significative property shared by all the multishell AGP approximations discussed above is their fast convergence rate as a function of the number of configurations. As seen from Table 3, the estimates of $s_2^{(or)}$ obtained after averaging the marginals over 5×10^3 configurations do in fact coincide with those produced with 10^4 configurations to the first decimal place.

B. TIP4P Water Close to the Temperature of Maximum Density. Upon lowering the temperature while keeping the pressure fixed at 1 bar, TIP4P water first exhibits the well-known maximum density anomaly at $T_{TMD} = 253 \pm 5$ K before congealing into ice I_h at $T_f = 232 \pm 5$ K, i.e., about 40 K below the experimental freezing point.[25] At 260 K, the average values of the specific density and excess internal energy were found to be 1.001 g cm^{-3} and − 10.67 kcal mol^{-1}, respectively. At lower temperatures, both the positional and angular order are more enhanced and longer-ranged. The RDF of the liquid is definitely more structured than at ambient conditions (see Figure 1), and we consistently found a value for the translational pair entropy (− 3.59 e.u.) that is about 18% lower than that obtained at 300 K. As for the calculation of the orientational pair entropy, we verified that the fit of $\sigma(r; N_{conf})$ was of comparable accuracy to that achieved at higher temperatures. Also in this case, the tail of $\sigma(r; N_{conf})$ turned out to scale as N_{conf}^{-1} at large distances. Figure 9 shows the extrapolated OLE and the corresponding integrand function. The comparison with the approximate estimates obtained from the four AGP schemes that we have already illustrated in the preceding sections confirms that even at this lower temperature the MSAGP1 approximation more faithfully reproduces the profile of $S^{(or)}(r)$, both at short and large distances. Correspondingly, the MSAGP1 estimate of the orientational pair entropy was again found to be closer than the other three approximate estimates to the asymptotic simulation value, the relative discrepancy being about 7% (see Table 2). We observe that $s_2^{(or)}$—and, correspondingly, the amount of angular order in water—is more significantly affected than $s_2^{(tr)}$ by the 40 K temperature drop. In fact, the value of the orientational pair entropy at 260 K was found to be about 29% lower than that at 300 K.

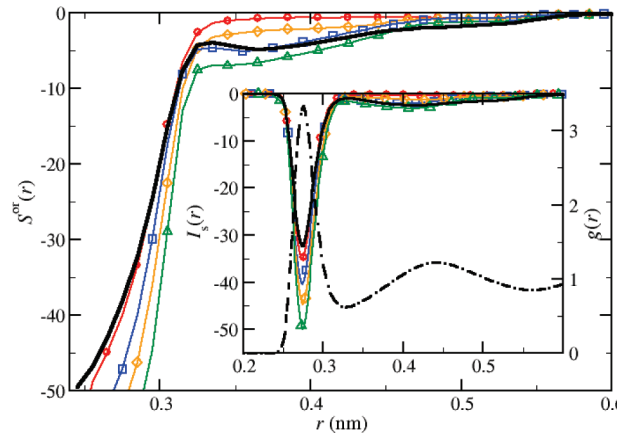

Figure 9. Orientational local entropy at $P = 1$ bar and $T = 260$ K: black curve, present asymptotic estimate obtained in the B_s set; red curve marked with circles, MSAGP estimate; blue curve marked with squares, MSAGP1 estimate; orange curve marked with diamonds, MSAGP2 estimate; green curve marked with triangles, MSAGP3 estimate. The inset shows the corresponding integrand functions $I_S^{(or)}(r)$ and the radial distribution function sampled over 10^5 configurations (right axis).

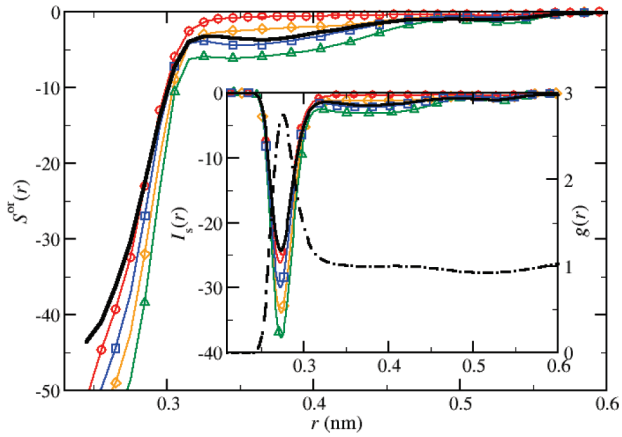

Figure 10. Orientational local entropy at $P = 4$ kbar and $T = 300$ K: black curve, present asymptotic estimate obtained in the B_s set; red curve marked with circles, MSAGP estimate; blue curve marked with squares, MSAGP1 estimate; orange curve marked with diamonds, MSAGP2 estimate; green curve marked with triangles, MSAGP3 estimate. The inset shows the corresponding integrand functions $I_S^{(or)}(r)$ and the radial distribution function sampled over 10^5 configurations (right axis).

C. TIP4P Water at Higher Pressure. We finally investigated the properties of ambient temperature TIP4P water compressed at a higher pressure ($P = 4$ kbar), falling in the range where crystalline ice II and ice III phases become stable at lower temperatures.[11,12] We found 1.134 g cm^{-3} and $-$ 10.18 kcal mol^{-1} for the average values of the specific density and excess internal energy, respectively. At variance with the behavior ordinarily observed in simple atomic fluids, the compression largely disrupts the local order observed in water at lower pressures. The effect on positional correlations is manifest in Figure 1. Notwithstanding this "antagonist" role played by the pressure, we found that, upon increasing P from 1 bar to 4 kbar at 300 K, the translational pair entropy dropped from -3.05 e.u. to -3.35 e.u.; the disruption of a relatively ordered network, which would imply a higher entropy, is more than compensated in this case by the reduction of available positional states produced by the 15% increase of the specific density. The decorrelating effect produced by compression is even stronger on angular order, as witnessed by the moderate increase that was registered instead in the orientational pair entropy (see Table 2), notwithstanding the increase of the density.[14]

The fit of $\sigma(r; N_{conf})$ turned out to be as accurate as that accomplished in the other two thermodynamic states. The scaling of the function as N_{conf}^{-1} at large distances was also confirmed. Figure 10 shows the extrapolated OLE and the corresponding integrand function. The comparison between the results obtained from the four AGP schemes confirms once more that even at higher pressures the MSAGP1 approximation is the most reliable approximation at short as well as large distances and also provides the most accurate estimate of the orientational pair entropy.

D. Comparison with Experimental Data. Table 4 presents a comparison between the cumulative pair entropies obtained from the current MD simulations without resorting to any approximation and the excess entropies of both TIP4P and ordinary water. The TIP4P values[26] of S_{ex} were obtained using thermodynamic integration methods for the calculation of the free energy, as extensively discussed in ref 12, while the experimental values follow from the data for the absolute entropy tabulated in ref 27, after subtracting the ideal-gas entropy. This latter contribution was calculated from eq (6.5) of ref 27, which parametrizes the ideal Helmholtz free energy. Note that the experimental estimate of the excess entropy at 260 K obviously refers to metastable undercooled water and was obtained upon extrapolating the values of the specific density and of the absolute entropy below the freezing temperature. We first observe that the pair entropy decreases upon lowering the temperature. The effect produced by an increase of the pressure on the local order of the liquid is more subtle, as already discussed above and more systematically analyzed in ref 14. In this specific instance, the effect is almost null, since the difference of 0.1 e.u. is presumably within the numerical uncertainty of the calculation. It should further be noted that the almost equal values found for S_2 at 300 K across the 4 kbar pressure gap is the outcome of differing relative weights of the translational and orientational pair entropies.

Table 4. Pair and Excess Entropies

		entropies (e.u.)		
T (K)	P (bar)	S_2 [TIP4P]	S_{ex} [TIP4P][a]	S_{ex} [expt.][b]
260	1×10^0	-18.39	-17.00	-15.73[c]
300	1×10^0	-14.57	-14.61	-13.99
300	4×10^3	-14.68	-15.14	-14.38

[a] Data from ref 26, reported with an estimated uncertainty of 0.06 e.u. [b] Estimates obtained from the data available in ref 27. [c] Estimate obtained upon extrapolating the properties of liquid water below the freezing point.

As shown in Table 4, at 260 K, the pair entropy actually overcomes the excess entropy: this implies a *positive* value of the so-called "residual multiparticle entropy" (RMPE), a quantity defined as the difference between S_{ex} and S_2.[19] As diffusely documented in the literature, a positive RMPE is evidence of a highly structured liquid.[13,28] On the other hand, at 300 K, the RMPE of TIP4P water almost vanishes at ambient pressure conditions, while being negative at 4 kbar but less than 3% of the total excess entropy. We remark that, as previously noted by Lazaridis and Karplus,[9] a small value of the RMPE does not necessarily imply that triplet or higher-order correlations do not play a role in determining the microscopic structure of the liquid. In fact, their overall contribution to the configurational entropy of a given substance may well be small or may even sum up to zero in some thermodynamic points or regions of the phase diagram, despite the fact that distribution functions beyond the pair one do not trivially reduce to the mere product of lower-order distribution functions.

It appears from Table 4 that at 300 K the pair entropy of TIP4P water provides fairly good estimates of the excess entropy of ordinary water. However, we are also aware that the agreement between the model and experimental data might be partially biased by the 40 K "shift" toward lower temperatures of the phase diagram predicted by the TIP4P model of water relatively to that of ordinary water.

V. Concluding Remarks

In this paper, we have presented a molecular dynamics calculation of the pair entropy of liquid water, modeled with the four-point transferable intermolecular potential (TIP4P) at three distinct thermodynamic states, corresponding to different values of temperature and pressure. The pair entropy is an integrated measure of two-body density correlations and represents the predominant contribution to the configurational entropy of a liquid. As such, it can be confidently used as a local structural estimate of the total configurational entropy: local, in that it does not call for an integration of the properties of the liquid along a thermodynamic path; structural, since it provides a direct connection between entropy and spatial order as monitored by the pair distribution function. This is the first aspect that we have put under scrutiny in this paper by extending preexisting analyses carried out for TIP4P water to thermodynamic states other than ambient conditions. Our new results for the orientational pair entropy follow from the calculation of the five-dimensional histogram that we obtained upon sampling, at given temperature and pressure, the configurations corresponding to different relative orientations of a generic water molecule with respect to a reference one, while keeping their centers of mass at a fixed distance. An intrinsic bias on the present results comes from the histogram bin width, with particular regard to the angular resolution. Our choice (10°), while being that commonly made in previous works on this subject, arises from a compromise between resolution and statistical quality of the calculations, a compromise that is unavoidably forced by the need for maintaining the overall size of the computation at a feasible level.

A secondary goal of this paper was that of discussing and testing some approximate schemes for the orientational distribution function that are based on the calculation of lower-order marginals, given the heavy computational task one has to face in a full-size calculation. In this respect, we have analyzed the performance of a number of factorizations that partially modify the "adjusted gas phase approximation" originally proposed by Lazaridis and Karplus.[9] We found that the best results, as far as both the orientational local entropy and its integrated value are concerned, were obtained at all of the three sampled thermodynamic states when using the so-called MSAGP1 approximation, which includes intramolecular correlations between the angle formed by the dipole vector of a water molecule and the intermolecular axis, and the angle describing the rotation of the same molecule about its own dipole vector. We emphasize that the MSAGP1 estimates were obtained upon sampling the marginal distribution functions over 10^4 configurations only and appear to underestimate the corresponding "exact" molecular dynamics values—which were obtained with a sampling carried out over a 10 times larger number of configurations—by 12% at the most.

On the other hand, we have also verified that including the intermolecular contribution which arises from cross correlations between the same angles mentioned above but measured on different molecules does actually worsen the agreement with the molecular dynamics results in that the ensuing modified scheme (MSAGP3) manifestly overestimates the degree of orientational order present in the liquid in a systematic way.

Acknowledgment. E.G. acknowledges a Ph.D. grant from CNISM (Italy) and the scientific hospitality of the Université Pierre et Marie Curie (Paris, France) in 2008 during a six-month visit cofunded by the Erasmus/Socrates program and also thanks Drs. Rubens Esposito and Dino Costa for numerical assistance at different stages of this work. P.V.G. thanks Prof. Carlos Vega for providing some reference thermodynamic data on the TIP4P model and Dr. Lingle Wang for some clarifying comments on the calculation of the orientational entropies with the *k*th nearest-neighbor method. The authors gratefully acknowledge the allocation of computer time from the French National Supercomputing Facility IDRIS, within the projects CP9-81387 and CP9-91387, and from the "Centro di Calcolo Elettronico "Attilio Villari" della Università degli Studi di Messina".

References

(1) Ben-Naim, A. *Molecular Theory of Water and Aqueous Solutions - Part 1: Understanding Water*; World Scientific: Singapore, 2009.

(2) Ball, P. *Nature* **2008**, *452*, 291.

(3) Green, H. S. *The Molecular Theory of Fluids*; North-Holland: Amsterdam, The Netherlands, 1952.

(4) Stratonovich, R. L. *Sov. Phys. JETP* **1955**, *28*, 409.

(5) Nettleton, R. E.; Green, M. S. *J. Chem. Phys.* **1958**, *29*, 1365.

(6) Yvon, J. *Correlations and Entropy in Classical Statistical Mechanics*; Pergamon Press: Oxford, United Kingdom, 1969.

(7) Baranyai, A.; Evans, D. J. *Phys. Rev. A* **1989**, *40*, 3817.

(8) Prestipino, S.; Giaquinta, P. V. *J. Stat. Phys.* **1999**, *96*, 135; **2000**, *98*, 507.

(9) Lazaridis, T.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 4294.

(10) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(11) Sanz, E.; Vega, C.; Abascal, J. L. F.; MacDowell, L. G. *J. Chem. Phys.* **2004**, *121*, 1165.

(12) Vega, C.; Sanz, E.; Abascal, J. L. F.; Noya, E. G. *J. Phys.: Condens. Matter* **2008**, *2*, 153101.

(13) Saija, F.; Saitta, A. M.; Giaquinta, P. V. *J. Chem. Phys.* **2003**, *119*, 3587.

(14) Esposito, R.; Saija, F.; Saitta, A. M.; Giaquinta, P. V. *Phys. Rev. E* **2006**, *73*, 040502(R).

(15) Zielkiewicz, J. *J. Chem. Phys.* **2005**, *123*, 104501.

(16) Zielkiewicz, J. *J. Chem. Phys.* **2006**, *124*, 109901.

(17) Zielkiewicz, J. *J. Phys. Chem. B* **2008**, *112*, 7810.

(18) Wang, L.; Abel, R.; Friesner, R. A.; Berne, B. J. *J. Chem. Theory Comput.* **2009**, *5*, 1462.

(19) Giaquinta, P. V.; Giunta, G. *Physica A* **1992**, *187*, 145.

(20) Giaquinta, P. V.; Giunta, G.; Prestipino Giarritta, S. *Phys. Rev. A* **1992**, *45*, R6966.

(21) Jakse, N.; Charpentier, I. *Phys. Rev. E* **2003**, *67*, 061203.

(22) Prestipino, S.; Giaquinta, P. V. *J. Stat. Mech.* **2004**, P09008.

(23) Martyna, G. J.; Tuckerman, M. E.; Klein, M. L. PINY_MD (c) Simulation Package, 2002.

(24) *IUPAC Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"). XML on-line corrected version: http://goldbook.iupac.org/E02151.html (accessed Nov 23,2009).

(25) Vega, C.; Abascal, J. L. F. *J. Chem. Phys.* **2005**, *123*, 144504.

(26) Vega, C. 2009, private communication.

(27) Wagner, W.; Pruss, A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387.

(28) Giaquinta, P. V. Entropy revisited: The interplay between entropy and correlations. In *Highlights in the Quantum Theory of Condensed Matter*; Beltram, F., Ed.; Publications of the Scuola Normale Superiore (Selections), Edizioni della Normale: Pisa, Italy, 2005; Vol. 1, pp 9−14. See also the introduction and the bibliography in Giaquinta, P. V. *Entropy* **2008**, *10*, 248.

CT900627Q