RESEARCH ARTICLE | SEPTEMBER 01 2023

A statistical mechanical model of supercooled water based on minimal clusters of correlated molecules

Isabella Daidone ; Riccardo Foffi ; Andrea Amadei ; Laura Zanetti-Polzi

J. Chem. Phys. 159, 094502 (2023)
https://doi.org/10.1063/5.0157505

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

A statistical mechanical model of supercooled water based on minimal clusters of correlated molecules

Cite as: J. Chem. Phys. 159, 094502 (2023); doi: 10.1063/5.0157505
Submitted: 8 May 2023 • Accepted: 14 August 2023 •
Published Online: 1 September 2023

Isabella Daidone,[1] Riccardo Foffi,[2] Andrea Amadei,[3,a)] and Laura Zanetti-Polzi[4,a)]

AFFILIATIONS

[1]Department of Physical and Chemical Sciences, University of L'Aquila, via Vetoio (Coppito 1), L'Aquila 67010, Italy
[2]Department of Civil, Environmental and Geomatic Engineering, Institute of Environmental Engineering, ETH Zürich, Laura-Hezner-Weg 7, 8093 Zürich, Switzerland
[3]Department of Chemical and Technological Sciences, University of Rome "Tor Vergata," Via della Ricerca Scientifica, I-00185 Rome, Italy
[4]Center S3, CNR-Institute of Nanoscience, Via Campi 213/A, 41125 Modena, Italy

[a)]**Authors to whom correspondence should be addressed:** andrea.amadei@uniroma2.it and laura.zanettipolzi@nano.cnr.it

ABSTRACT

In this paper, we apply a theoretical model for fluid state thermodynamics to investigate simulated water in supercooled conditions. This model, which we recently proposed and applied to sub- and super-critical fluid water [Zanetti-Polzi *et al.*, J. Chem. Phys. **156**(4), 44506 (2022)], is based on a combination of the moment-generating functions of the enthalpy and volume fluctuations as provided by two gamma distributions and provides the free energy of the system as well as other relevant thermodynamic quantities. The application we make here provides a thermodynamic description of supercooled water fully consistent with that expected by crossing the liquid–liquid Widom line, indicating the presence of two distinct liquid states. In particular, the present model accurately reproduces the Widom line temperatures estimated with other two-state models and well describes the heat capacity anomalies. Differently from previous models, according to our description, a cluster of molecules that extends beyond the first hydration shell is necessary to discriminate between the statistical fluctuation regimes typical of the two liquid states.

I. INTRODUCTION

The anomalous thermodynamic properties of water become more pronounced in deeply supercooled conditions.[1] Among other theoretical scenarios to explain water anomalies, the existence of a liquid–liquid phase transition (LLPT), and corresponding liquid–liquid critical point (LLCP), has been hypothesized in the supercooled region.[2] According to this scenario, a first order phase transition exists between two liquid states differing in density: the high density liquid (HDL) and the low density liquid (LDL). The extension of the coexistence line associated with the phase transition beyond the second critical point, the Widom line, is typically considered as providing the border between regions mostly characterized by either the HDL-like or the LDL-like phase. The LLPT and LLCP are supposed to lie in the so called no man's land, where experimental observation is hampered by fast crystallization. Therefore, the investigation of supercooled water, and its thermodynamics, received strong support from theoretical-computational approaches.[3–7]

A remarkably successful and versatile approach to study the thermodynamics of the liquid–liquid transition can be found in so-called two-state (or binary mixture) models.[1] The fundamental idea behind such models is the existence of two distinct but continuously interconverting local structures, each of which is representative of one of the two phases of water: LDL-like structures with lower energy and entropy, and HDL-like structures where energy and entropy are higher. The competition between these two local structures can lead to phase separation if the non-ideal component of the Gibbs free energy of mixing is strong enough. The origin of the non-ideality in this mixing term remains an interesting point

of discussion. While for models such as ST2, the transition was shown to be energy-driven,[8] criticality in real water was proposed to be mainly entropy-driven.[9] Recently, a minimal microscopic model that does not rely on phenomenological assumptions was developed by Caupin and Anisimov,[10] providing a unified framework to describe polymorphic liquids. Many researchers have instead focused on connecting the two-state model to explicit microscopic structural descriptions, defining structural quantities that could serve as order parameters in these equations of state.[11–18] A structural order parameter-free approach to distinguish the two states on the basis of their entropic and energetic differences has also recently been proposed.[19] Despite numerous efforts, a complete microscopic description of water's local structures has yet to be achieved. Foffi and Sciortino[18] have shown that, close to the liquid–liquid critical point, most of the existing structural descriptors accurately follow the critical density fluctuations, sampling a clearly bimodal free energy landscape; in parallel, work by Offei-Danso *et al.*[20] based on a new unsupervised learning approach has shown the impossibility of associating individual molecules to distinct local structures in ambient-condition water through standard structural descriptors.

Recently,[21] we proposed a statistical mechanical model to describe the thermodynamics of fluids based on the quasi-Gaussian entropy (QGE) theory.[22,23] In this model, the theoretical background of which is briefly recalled in Sec. II, we used the distributions of the fluctuations of macroscopic properties to describe the thermodynamics of the system as a function of both temperature and pressure deriving a complete equation of state in the isothermal–isobaric ensemble.

In our previous work,[21] we applied this model to the liquid to gas transitions in sub- and super-critical fluid water. We showed that fluid water thermodynamics can be well described using two gamma state solutions (i.e., solutions based on Gamma distributions): a liquid-like gamma state and a gas-like gamma state, able to describe, respectively, the liquid/pseudo–liquid and gas/pseudo–gas phases in both the sub- and super-critical p, T spaces. We also modeled the phase or pseudo-phase transitions by introducing a new parameter, namely N_c, which represents the minimal size of the cluster of correlated molecules inside the macroscopic system. In other words, interpreting the macroscopic system as a set of subsystems in either the liquid-like or gas-like state, N_c corresponds to the size of the smallest subsystem that can be found in the liquid-like and gas-like states. According to this interpretation, in the sub-critical region $N_c \to \infty$, providing the discontinuous behavior of macroscopic phase transitions. From the comparison between experimental water data and our equation of state, we obtained in the super-critical region $N_c = 5$, i.e., the smallest subsystem size in the supercritical fluid corresponds to the number of molecules involved in the minimal hydrogen-bonding (HB) network. In our model, the macroscopic phase transition is only determined by the $N_c \to \infty$ limit, as the non-ideal term in the free energy of mixing vanishes within the assumptions/approximations used.[21]

In the present work, we use the same model to investigate the properties of supercooled water in the region where a continuous transition between the two liquid phases differing in density is expected by crossing the liquid–liquid Widom line. To achieve this aim, we perform molecular dynamics (MD) simulations of the TIP4P/2005 water model, which was shown several times to display a LLPT,[24–26] using various temperatures along different isobars. We apply the above-introduced model along these isobars without explicitly including the pressure dependence, obtaining, therefore, the thermodynamic properties as a function of the temperature at each pressure. In analogy with what we did for the liquid to gas phase transition, we use two gamma state solutions to model the LDL-like and HDL-like phases. Using the enthalpy fluctuation moment generating functions, as provided by the two gamma distributions, we gain insights into the thermodynamic properties, such as the Gibbs free energy (G) and heat capacity (C_p), of the LDL-like and HDL-like states as a function of the temperature. In addition, we also discuss the behavior of the parameter N_c along the investigated isobars. Finally, we compare the results of our model with those provided by the use of well established order parameters that make use of specific structural features to investigate the LDL-like and HDL-like phases.

II. THEORY

In the QGE theory, the basic statistical mechanical relations are rewritten in terms of the distributions of the fluctuations of macroscopic properties. Thanks to that, it is possible to describe the thermodynamics of the system without evaluating the partition function. It was previously shown[22,23,27] that the thermodynamics of fluid state systems can be well described by making use of Gamma distributions to model the fluctuation distributions of thermodynamic properties (such as enthalpy and volume). The use of Gamma distributions permits the analytical derivation of relevant thermodynamic quantities avoiding the use of empirical functions or parameters, providing a physically coherent description of the system under investigation. Below, we provide the main equations used in the present paper to describe the isobaric thermodynamics of water in the supercooled region. The theoretical basis of the approach and the full derivations can be found in the original work.[21]

Given a fluid system of N molecules in the isobaric–isothermal (N, p, T) ensemble, its Gibbs free energy is given by

$$G(p,T) = -k_B T \ln \Delta(p,T) \tag{1}$$

with k_B is the Boltzmann constant and T is the absolute temperature, and $\Delta(p,T)$ is the isobaric–isothermal partition function,

$$\Delta(p,T) = \sum_V Q(V,T) e^{-\beta p V}, \tag{2}$$

$$Q(V,T) = \Theta \sum_l \int_V e^{-\beta[\mathscr{U}_e(\boldsymbol{q}) + \mathscr{E}_{vb,l}(\boldsymbol{q}) + \mathscr{K}(\boldsymbol{q},\boldsymbol{\pi})]} d\boldsymbol{\Gamma}. \tag{3}$$

In Eqs. (2) and (3), p is the equilibrium pressure, $1/\beta = k_B T$, the summation in Eq. (2) is over all the possible volumes V of the system (the difference between two consecutive volumes is virtually corresponding to a differential), the summation in the canonical partition function $Q(V,T)$ is over all the vibrational states, and the subscript V of the integral sign means that integration is performed within the volume V. Moreover, $\mathscr{U}_e(\boldsymbol{q})$ is the electronic ground state energy (the electronic excited states are disregarded as they are virtually inaccessible except at extremely high temperatures), $\mathscr{E}_{vb,l}(\boldsymbol{q})$ is the lth vibrational state energy, $\mathscr{K}(\boldsymbol{q},\boldsymbol{\pi})$ is the classical kinetic energy, Θ is a constant providing the quantum corrections for the permutations of identical particles (possibly including the degeneration

factor of the electronic ground state), and $d\mathbf{\Gamma} = d\boldsymbol{q}d\boldsymbol{\pi}/h^n$ expresses the number of semiclassical quantum states within the phase space differential with h being the Planck's constant, $\boldsymbol{q}$ being the n generalized semiclassical coordinates, and $\boldsymbol{\pi}$ being the corresponding conjugated momenta. By expanding the exponential term in Eq. (3), Eq. (1) can be rewritten as follows (see Zanetti-Polzi *et al.*[21]):

$$\beta G(p,T) \cong -\ln Q_{vb}(T) - \ln \sum_V \Theta \int_V e^{-\beta[\mathcal{U}_e(\boldsymbol{q})+\mathcal{K}(\boldsymbol{q},\boldsymbol{\pi})+pV]} d\mathbf{\Gamma} \tag{4}$$

with

$$Q_{vb}(T) = \sum_l e^{-\beta E_{vb,l}}, \tag{5}$$

the vibrational partition function, where $E_{vb,l}$ is the lth vibrational state mean energy.[21]

In a statistical mechanical ensemble, it is always possible to define a proper reference state such that the free energy difference between the actual condition and the reference one can be expressed in terms of the moment-generating function of the distribution of a specific macroscopic fluctuation. Setting at T_0 the temperature of this reference state and defining the excess Gibbs free energy as

$$G' \cong G + k_B T \ln Q_{vb}. \tag{6}$$

It can be shown[21] that the excess free energy change can be expressed using the moment generating function of the distribution of the single phase space position enthalpy $\mathcal{U}_e(\boldsymbol{q}) + \mathcal{K}(\boldsymbol{q},\boldsymbol{\pi}) + pV$,

$$\beta G'(p,T) - \beta_0 G'(p,T_0) = -\ln \langle e^{-\Delta\beta[\mathcal{U}_e(\boldsymbol{q})+\mathcal{K}(\boldsymbol{q},\boldsymbol{\pi})+pV]} \rangle_{\beta_0}, \tag{7}$$

where the β_0 subscript of the angle brackets means averaging within the β_0 ensemble. Using for such distribution the diverging Gamma distribution and its moment generating function[21] and given the excess entropy, enthalpy, and heat capacity (s'_0, h'_0, and c'_{p0}) at the arbitrary reference temperature T_0, we can derive the excess thermodynamics along an isobar according to the diverging gamma state model

$$\mu'(p,T) = \frac{G'(p,T)}{N} = h'_0 - T_0 c'_{p0} + T(c'_{p0} - s'_0) + T c'_{p0} \ln \frac{T_0}{T} = \mu'_0 + (T - T_0)(c'_{p0} - s'_0) + T c'_{p0} \ln \frac{T_0}{T}, \tag{8}$$

$$s'(p,T) = \frac{S'(p,T)}{N} = s'_0 + c'_{p0} \ln \frac{T}{T_0}, \tag{9}$$

$$h'(p,T) = \frac{H'(p,T)}{N} = h'_0 + c'_{p0}(T - T_0), \tag{10}$$

$$c'(p,T) = \frac{C'_p(p,T)}{N} = c'_{p0}, \tag{11}$$

Notably, the linear temperature dependence obtained for the excess enthalpy $h'(p,T)$ can be used as the diagnostic criterion for validating the diverging Gamma state as a proper model for the isobar thermodynamics. By using Eqs. (6) and (8), we can obtain the complete molecular Gibbs free energy (i.e., the chemical potential) along the isobar

$$\mu(p,T) \cong \mu'(p,T) - k_B T \ln q_{vb}(T) = \mu'_0 + (T - T_0)(c'_{p0} - s'_0) + T c'_{p0} \ln \frac{T_0}{T} - k_B T \ln q_{vb}(T) \tag{12}$$

with q_{vb} being the molecular vibrational partition function and furnishing via its temperature derivatives the corresponding molecular full entropy, enthalpy, and (isobaric) heat capacity.

If we can conceive the macroscopic system as constituted by a mixture of M subsystems, each including N_c molecules and corresponding to either the LDL-like or the HDL-like state, we can construct a general model for the LDL ⇋ HDL equilibrium. In fact, defining with μ^*_{LDL} and μ^*_{HDL} the LDL and HDL chemical potentials when the whole macroscopic system is either in the LDL or HDL state, respectively, the total free energy of the macroscopic system (for a given subsystem LDL-HDL configuration) can be expressed by[21]

$$G_{conf} = \sum_{j=1}^{M} N_c \mu_{conf,j} \tag{13}$$

with $\mu_{conf,j}$ being the chemical potential of the molecules within the jth subsystem, corresponding to either the LDL ($\mu_{conf,L}$) or HDL ($\mu_{conf,H}$) one,

$$\mu_{conf,L} \cong \mu^*_L + \chi_H \delta_L, \tag{14}$$

$$\mu_{conf,H} \cong \mu^*_H + \chi_L \delta_H = \mu^*_H + (1 - \chi_H)\delta_H, \tag{15}$$

with χ_H being the fraction of the M subsystems in the HDL-like state, $\chi_L = 1 - \chi_H$ being the fraction of the LDL-like subsystems, and δ_L, δ_H being the LDL-like and HDL-like chemical potential corrections when considering either a single LDL-like subsystem into a HDL-like macroscopic system (i.e., $\chi_H \to 1$) or a single HDL-like subsystem into a LDL-like macroscopic system (i.e., $\chi_L \to 1$). Note that in the last equations, the terms providing inhomogeneity effects neglect any *local* effect due to the neighboring subsystems, assuming the free energy corrections only depend on the overall fractions χ_H, χ_L. It is also worth noting that these free energy corrections possibly include surface effects due to such inhomogeneities. However, when the subsystem molecular size (i.e., N_c) is as small as in the present case (10–20 molecules, see below), there is no phase separation; thus, surface tension can be disregarded.

Using Eqs. (13)–(15) and assuming $\delta_L \cong -\delta_H$, we obtain (considering all the subsystem configurations providing all the combinations of k LDL-like and $M-k$ HDL-like subsystems) the isothermal–isobaric partition function and corresponding chemical potential[21]

$$\Delta(p,T) \cong \sum_{k=0}^{M} \frac{M!}{k!(M-k)!} e^{-\beta G(k)} \cong \sum_{k=0}^{M} \frac{M!}{k!(M-k)!} e^{-\beta k N_c \mu^*_L} e^{-\beta(M-k)N_c \mu^*_H} = \left(e^{-\beta N_c \mu^*_L} + e^{-\beta N_c \mu^*_H}\right)^M, \tag{16}$$

$$\mu(p,T) = \frac{G(p,T)}{N} = -\frac{k_B T}{N}\ln\Delta(p,T) \cong -\frac{k_B T}{N_c}\ln\left(e^{-\beta N_c \mu_L^*} + e^{-\beta N_c \mu_H^*}\right) \quad (17)$$

with $N = N_c M$ being the total number of water molecules in the macroscopic system. The approximation used $\delta_L \cong -\delta_H$ is based on the fact that a molecule within an HDL subsystem embedded in LDL subsystems experiences an increase in the average number of hydrogen bonds roughly equal to the decrease in the average number of hydrogen bonds experienced by a molecule within a LDL subsystem embedded in HDL subsystems. From the above-mentioned equation, we can obtain the molecular enthalpy, entropy, and heat capacity, as follows:

$$h(p,T) = \left(\frac{\partial \beta\mu}{\partial\beta}\right)_p \cong \frac{e^{-\beta N_c\mu_L^*} h_L^*}{e^{-\beta N_c\mu_L^*} + e^{-\beta N_c\mu_H^*}} + \frac{e^{-\beta N_c\mu_H^*} h_H^*}{e^{-\beta N_c\mu_L^*} + e^{-\beta N_c\mu_H^*}}, \quad (18)$$

$$s(p,T) = -\left(\frac{\partial \mu}{\partial T}\right)_p \cong \frac{e^{-\beta N_c\mu_L^*} s_L^*}{e^{-\beta N_c\mu_L^*} + e^{-\beta N_c\mu_H^*}} + \frac{e^{-\beta N_c\mu_H^*} s_H^*}{e^{-\beta N_c\mu_L^*} + e^{-\beta N_c\mu_H^*}}, \quad (19)$$

$$c_p(p,T) = \left(\frac{\partial h}{\partial T}\right)_p \cong c_{p,L}^* + \frac{e^{-\beta N_c\Delta\mu^*}\,\Delta c_p^*}{1+e^{-\beta N_c\Delta\mu^*}} + \frac{N_c(\Delta h^*)^2\, e^{-\beta N_c\Delta\mu^*}}{k_B T^2\left(1+e^{-\beta N_c\Delta\mu^*}\right)^2}, \quad (20)$$

with $\Delta\mu^* = \mu_H^* - \mu_L^*$, $\Delta h^* = h_H^* - h_L^*$, and $\Delta c_p^* = c_{p,H}^* - c_{p,L}^*$, with the star superscript indicating that the property refers to the pure LDL or HDL macroscopic system.

We remark that here, unlike other two-state models, we do not deal with the interconversion of individual molecules between two distinct local structures; rather, our equation of state represents the interconversion between spatially extended clusters of molecules. A cluster-based approach to the two-state model was also adopted by Holten *et al.*[28] to describe the thermodynamics of mW water, obtaining an equation of state formally identical to that of a polymer solution in Flory–Huggins theory.[29] In that case, the cluster dimension N is a purely phenomenological parameter included to provide a reduction of the mixing entropy. In our model, instead, the N_c parameter follows from an explicit physical model based on constructing the thermodynamics on the cluster domains.

III. RESULTS AND DISCUSSION

To model the thermodynamics of supercooled water, we perform MD simulations of the TIP4P/2005 water model[30] along five different isobars (details on the MD simulations are provided in the supplementary material). As shown in Fig. 1(a), we choose along the first three isobars (i.e., 1, 400, and 1000 bars) a range of temperatures that crosses the previously predicted Widom line for the TIP4P/2005 water model.[24,25] Therefore, along these three isobars, a continuous

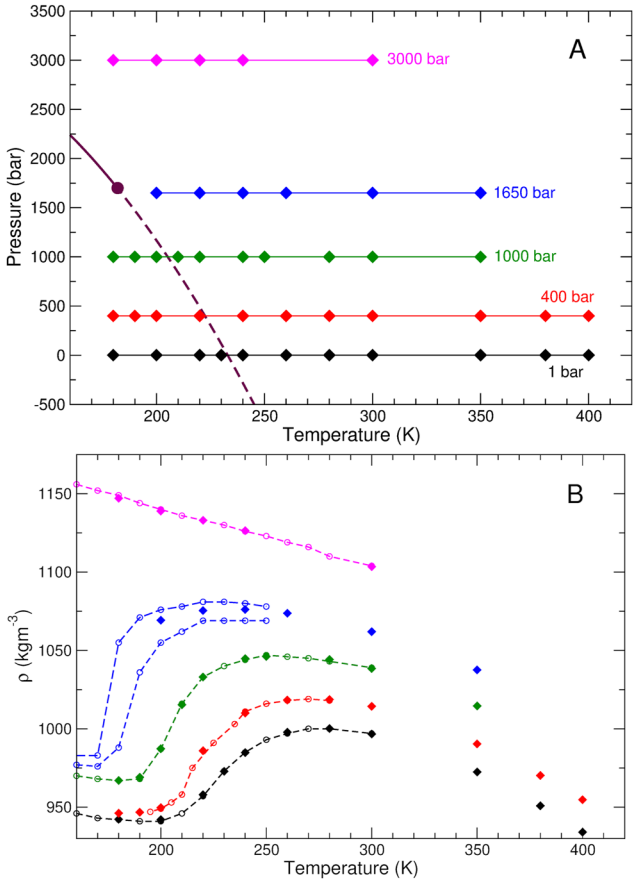

FIG. 1. (a): Filled diamonds identify the MD simulation temperatures along the five isobars: 1 bar (black), 400 bars (red), 1000 bars (green), 1650 bars (blue), and 3000 bars (magenta). In dark red, we report the liquid–liquid coexistence line (solid), Widom line (dashed), and LLCP (filled circle) as estimated by the two-structure equation of state (TSEOS) for the TIP4P/2005 water model by Singh *et al.*[24] (b): Filled diamonds correspond to the densities as obtained in our MD simulations along the five isobars: 1 bar (black), 400 bars (red), 1000 bars (green), 1650 bars (blue), and 3000 bars (magenta). Previously obtained[24] densities for the same water model are reported as open circles. Colored dashed lines are a guide for the eye. Since we could not find previous results at 1650 bars, we compare our simulated densities with those previously reported at 1500 bars (bottom blue dashed line) and 1750 bars (top blue dashed line).

transition between the LDL- and HDL-like phases can be observed. At 1650 and 3000 bars, instead, water is in the HDL-like phase at all the simulated temperatures. As a benchmark for our simulation conditions, we also compare, in Fig. 1(b), the densities as obtained from our MD simulations at all temperatures and pressures with those previously obtained for the same water model[24] (more details are provided in the supplementary material).

Using the average volume and energy as obtained by averaging over the MD simulation frames, we can then obtain the system excess enthalpy for each isobar as a function of the temperature. The collected trends, reported in Fig. 2, show that at 1650 and 3000 bars [panel (d), blue and magenta lines, respectively], the excess enthalpy shows a linear temperature dependence. Since along these

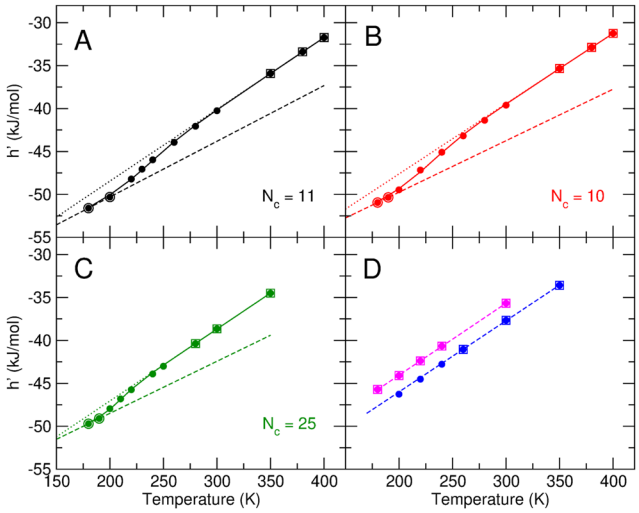

FIG. 2. Calculated excess enthalpy at 1 bar (a), 400 bars (b), 1000 bars (c), 1650 and 3000 bars [(d), blue and magenta, respectively]. At each pressure, the enthalpy values as obtained from the MD simulations are reported as filled circles, while the corresponding enthalpy trend provided by the model is reported as a solid line in panels (a) to (c). The points used to define the LDL- and HDL-like gamma states are highlighted as open squares and circles, respectively. The linear trends of the enthalpy corresponding to the full LDL- and HDL-like gamma states in the whole temperature range are reported as dashed and dotted lines, respectively.

two isobars, water is only found in the HDL-like state, the linearity of the enthalpy shows that the HDL-like state can be well modeled as a single gamma state. At 1650 bars, taking as reference for the HDL-like gamma state the points at the three highest temperatures [highlighted with blue squares in Fig. 2(d)], a slight deviation from the linearity is observed at the lowest temperature (i.e., 200 K), which is close to the estimated Widom line temperature (i.e., ≈185 K). This behavior is consistent with an isobar slightly below the critical pressure, for which a continuous yet sharp transition is expected by crossing the Widom line. The enthalpy trends at 1, 400, and 1000 bars [panels (a), (b), and (c), respectively] show instead an evident non-linear behavior. Therefore, liquid water at these pressures cannot be described by a unique gamma state. Yet, at the highest and lowest temperatures, linearity is recovered, suggesting that along these three isobars, an interconversion between two different statistical fluctuation regimes takes place and that each regime can be modeled as a gamma state. In what follows, we will refer to these two distinct gamma states as LDL-like (low temperature) and HDL-like (high temperature) gamma states. Nonetheless, it has to be noted that the definition of these two states as emerging from our model does not exactly overlap with the definition of LDL- and HDL-like molecules that can be obtained with order parameters based on structural features that have been developed to investigate the properties of supercooled liquid water.[12,14–17,31–33] Similarly, the two gamma states, defined on the basis of the system statistical fluctuations, are not expected to exactly overlap with the LDL- and HDL-like phases obtained by means of two state models that treat water as a binary mixture of two different local structures.[9–11,24,25]

To define the LDL- and HDL-like gamma states at each pressure, we make use of the lowest and highest simulated temperatures (highlighted with open circles and squares, respectively, in Fig. 2). Given the difficulties in obtaining reliable results from very low temperature MD simulations, only two points are used to define the LDL-like gamma state. The actual linear temperature dependence of the enthalpy, implying that the LDL-like gamma state is properly identified, can only be confirmed from the estimation of the Widom line temperature at each pressure (*vide infra*). The linearity of the enthalpy in the high temperature region can instead be well appreciated in Fig. 2. Interestingly, it can be observed that the linear region identifying the HDL-like phase comprises temperatures that are not used to define the gamma state, showing that at all the investigated pressures in TIP4P/2005 water is found in a single fully HDL-like state above ≈280 K.

Once the LDL- and HDL-like gamma states are defined at each pressure, by using Eqs. (8), (10), and (18), we can obtain the model enthalpy curve by fitting the enthalpy values obtained from the MD simulations (a summary of the fitting parameters is provided in Table 2 in the supplementary material). The fit provides N_c and T_{eq}, with T_{eq} the equilibrium temperature, defined as the temperature at which $\Delta\mu^* = 0$. Following the definition used by Singh *et al.* in the two-structure equation of state (TSEOS),[24] in this paper we choose to define the Widom line temperature, T_W, as coinciding with T_{eq}. The values obtained for N_c are reported in Fig. 2, while those for T_W are reported in Fig. 3(b) (open diamonds). The values obtained for N_c are of the same order of magnitude for all the investigated pressures. More specifically, we obtain very similar N_c values at 1 and 400 bars (11 and 10, respectively) and an increased N_c value at 1000 bars (25). Despite these values do not show a clear trend, we cannot exclude that N_c might slightly increase with the pressure and that this increase cannot be observed between 1 and 400 bars due to the noise of the simulated data. In particular, the enthalpy estimate from the low temperature MD simulations might be affected by inaccuracies resulting in slight N_c changes. On the other hand, the similarity among the three N_c values obtained from the fit suggests that, similarly to what observed for the water liquid–vapor transition, a unique N_c value at all pressures could be a proper choice to describe the liquid–liquid Widom line crossing. We recall indeed that in the super-critical region of the liquid–vapor phase transition, we found in our previous work[21] that using $N_c = 5$ at all temperatures and pressures, the experimental supercritical thermodynamics was well reproduced. In analogy to this result, and assuming that the liquid–gas and the liquid–liquid phase transitions should behave similarly, we speculate that a single N_c value could provide a proper description of the thermodynamics of the super-critical regime of the liquid–liquid phase transition. As a matter of fact, using as unique N_c the average among the three fitted values (i.e., 15), the enthalpy trends are still very well fitted (see Fig. 2 of the supplementary material) with negligible effects on the estimate of T_W (Fig. 3 of the supplementary material).

Recalling that N_c is the minimum number of water molecules necessary to define a cluster that can be found in one of the two gamma states, we note that to describe the existence of one liquid phase into the other, a larger number of molecules is necessary with respect to what observed for the liquid-like and gas-like phases. In that case, $N_c = 5$ corresponds to the minimal HB pattern around a central molecule, and each subsystem of N_c molecules can be pictured as fluctuating between forming and disrupting this HB network. In the super-critical region of the liquid–liquid phase tran-

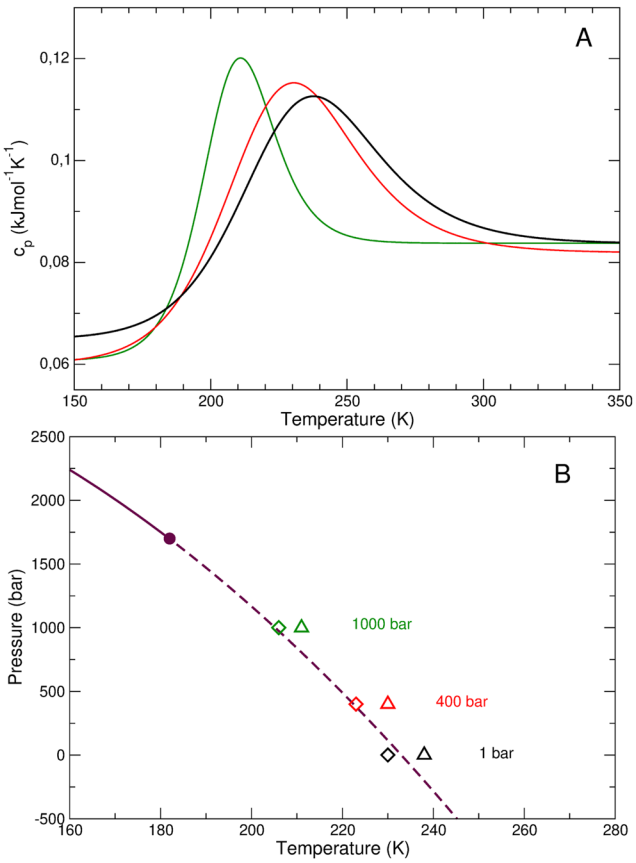

FIG. 3. (a): Isobaric heat capacity along the three isobars: 1 bar (black), 400 bars (red), and 1000 bars (green). (b): the Widom line temperature T_W obtained from the fits of the enthalpy trends reported in Fig. 2 is reported for the three pressures (open diamonds: 1 bar, black; 400 bars, red, and 1000 bars, green) and compared with that obtained with the TSEOS for TIP4P/2005 by Singh *et al.*[24] (dark red dashed line). The liquid–liquid coexistence line and LLCP as estimated by the TSEOS are also reported as the dark red solid line and filled circle, respectively. The c_p maxima obtained by our model are reported as open triangles (same color code for the three pressures).

sition, the minimal HB network is not sufficiently big to define one of the two liquid phases differing in density (in Fig. 4 of the supplementary material, we show that if we use N_c = 5, the data are poorly fitted). This observation agrees well with the fact that most of the difference between LDL- and HDL-like states was shown to depend on the second hydration shell, which is shifted toward lower radii upon raising the density.[34–37] As a matter of fact, N_c = 15 (i.e., the average among the three N_c values obtained from the fir at the three pressures) roughly corresponds to the number of water molecules up to the second hydration shell as provided by the O–O radial distribution function (see Fig. 5 of the supplementary material). Therefore, each subsystem can be pictured as fluctuating between the LDL-like state, in which the first 10–20 first neighbors are arranged around the central molecule in two well defined shells, and the HDL-like state, in which the boundary between the first and second hydration shells is less defined and the space between the two shells is filled by interstitial water molecules.

As shown in Fig. 3(b), the T_W values obtained from the fit are in good agreement with those obtained from the TSEOS by Singh *et al.* for the same water model.[24] In both models, T_W is defined as the temperature at which $\Delta\mu^* = 0$, showing that the two theoretical descriptions, although based on different physical concepts, provide a similar chemical potential behavior. In Fig. 3(a), we also report the isobaric heat capacity, c_p, as a function of the temperature as provided by the enthalpy trend obtained from the model [see Eq. (20) and Fig. 2]. In Fig. 6(A), in the supplementary material, we also report the comparison between the c_p trends obtained from the model and those obtained by the numerical temperature derivative of the enthalpies extracted from the MD data (i.e., the filled circles in Fig. 2). We note that, since the simulated enthalpies are very well reproduced by the model (see Fig. 2), the discrepancies between the model and the MD derived c_p trends can be ascribed to the limited temperature resolution in the numerical derivative. The overall c_p trend, both obtained by the model and obtained by the numerical temperature derivative, is in agreement with previous calculations under the same pressure conditions.[19] This is also true for the estimate we make of the isothermal compressibility, k_T, from our MD simulations [see Fig. 6(B) in the supplementary material], which are in agreement with previous corresponding calculations. In Fig. 3(b), the temperatures of the c_p maxima obtained at the three pressures are also reported (open triangles). We note that, according to our model, the c_p maxima are found at higher temperatures with respect to T_W by 7–8 K (with a slightly reduced temperature difference at 1000 bars, in agreement with the fact that at the critical point the Widom line and the loci of maxima of c_p coincide). We also note that recent estimates of the c_p maxima along the same isobars for the same water model provided somewhat lower temperatures for the c_p maxima (specifically, 220 vs 238 K; 212 vs 230 K; and 200 vs 211 K for isobars 1, 400, and 1000 bars, respectively). We cannot unequivocally determine the origin of the difference between the two estimates of the isobaric heat capacity temperature maxima; we note, however, that the heat capacity trend is very sensitive to small variations of the internal energy and density of the systems along the MD simulations, suggesting that the c_p is not the best observable to be compared among different models or calculation procedures.

From the temperature trend of $\Delta\mu$ provided by the model, we can also obtain the fraction of LDL- and HDL-like gamma states as a function of the temperature at each pressure. In Fig. 4, we report the LDL-like fraction at 1 and 1000 bars and compare it with the same quantity as provided by two different order parameters, namely ζ and ψ. The value ζ_i of a molecule i is defined as the difference between the distance $r_{j'i}$ of the first neighbor not hydrogen bonded to i and the distance $r_{j''i}$ of the last neighbor hydrogen bonded to i.[14] ψ is a recently introduced order parameter based on the observation that the distance between pairs of molecules separated by four links along the HB network (chemical distance D = 4) is different in the LDL- and HDL-like phases. The value ψ_i of a molecule i is defined as the minimal distance between molecules at chemical distance D = 4 from i.[18] To define the LDL-like fraction using the two order parameters, we use as a threshold to separate the two states the crossing point between the distributions obtained at the lowest and highest temperatures (see Fig. 7 in the supplementary material). It can be observed that, according to the model, the LDL-like fraction is

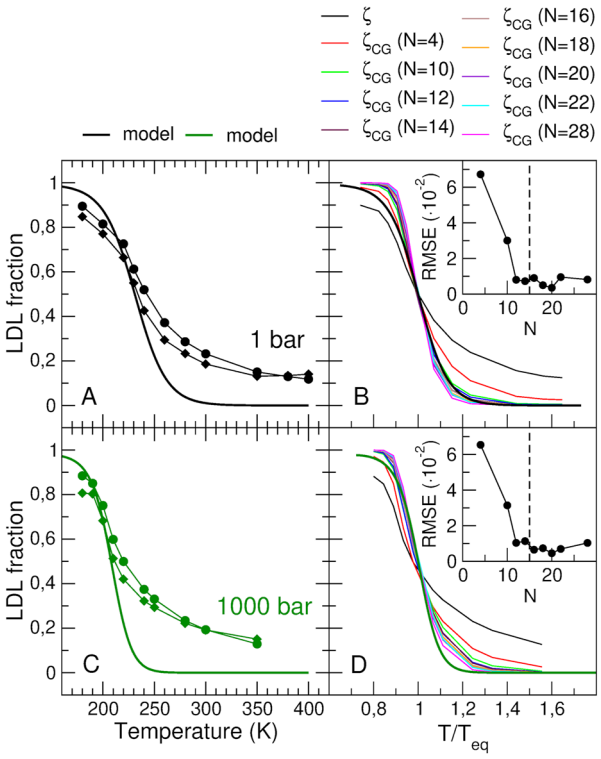

FIG. 4. (a) and (c): Fraction of LDL-like water as a function of the temperature at 1 bar (a) and 1000 bars (c) as provided by our model (solid line), by the ζ order parameter (filled circles) and by the ψ order parameter (filled diamonds). (b) and (d): Fraction of LDL-like water as a function of T/T_{eq}, with T_{eq} the temperature at which the model and ζ provide LDL fraction = 0.5, at 1 bar (b) and 1000 bars (d) as provided by the coarse grained version of the ζ order parameter using an increasing number N of nearest neighbors to perform the coarse-graining. Insets in panels (b) and (d) report the root mean square error (RMSE) between the curves by increasing N. The black dashed line highlights $N = 15$.

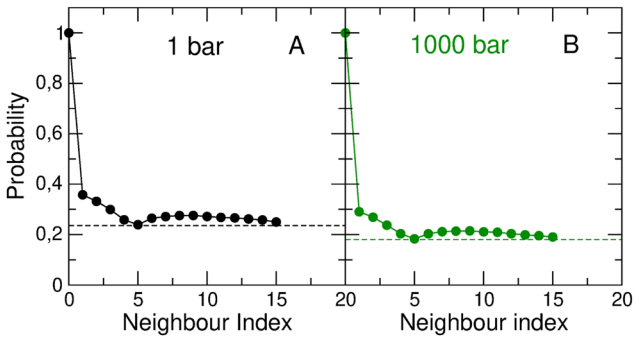

FIG. 5. Probability of being identified as an LDL-like molecule for the first N_c neighbors of a LDL-like molecule. Identification of the LDL-like molecules is done using the ζ order parameter. Panel (a): data at 1 bar and 300 K, $N_c = 15$. Panel (b): data at 1000 bars and 300 K, $N_c = 15$.

null for temperatures above ≈250–280 K (at 1000 and 1 bars, respectively), as anticipated by the linearity of the enthalpy trend at high temperatures. In the same temperature range, the order parameters provide instead a non zero LDL-like fraction (≈10%–15%), reaching at 1 bar a plateau value of ≈10% even at the highest temperature (400 K).

This apparent discrepancy can be rationalized by taking into account the interpretation of the N_c parameter in the model. As a matter of fact, N_c is the minimal cluster of water molecules that can be found in the LDL- or HDL-like gamma states and is, therefore, the smallest number of water molecules necessary to distinguish the two regimes of statistical fluctuations. While it is reasonable that at high temperatures some molecules have some structural features that can be identified as LDL-like (in terms of, e.g., tetrahedrality, absence of interstitial water molecules, or chemical distance as defined according to the ψ parameter), it is highly unlikely that a cluster of N_c nearest neighbors has these LDL-like features. This can be appreciated in Fig. 5, where we report the probability that we can find, using ζ as an order parameter, a LDL-like molecule among the first 15 nearest neighbors of an LDL-like molecule. It can be observed that at high temperatures, the probability that each of the first neighbors of an LDL-like molecule be also LDL is rather low. The probability of having all these neighbors identified as LDL-like is, therefore, negligible. Therefore, according to the ζ parameter, a cluster of 10–20 LDL-like water molecules is not found at high temperatures. Coherently with this observation, the LDL-like fraction at high temperatures obtained with our model is null. This also implies that single isolated water molecules that possess, at high temperatures, structural LDL-like features cannot be assigned to the LDL-like phase if a more extended structural organization is taken into account. Yet, according to our model, a cluster of molecules that extends beyond the first hydration shell is necessary to discriminate between the statistical fluctuation regimes typical of the two phases.

Recently, spatial coarse-graining was proposed as a strategy to take into account the effect of nearest neighbors in calculating the ζ order parameter.[38] The coarse-grained value of ζ, ζ_{CG}, for the molecules i is obtained by averaging its ζ value with that of the molecules within its first shell (i.e., within 0.35 nm). We show in Figs. 4(b) and 4(d) that by calculating ζ_{CG} including an increasing number N of nearest neighbors (going beyond the first shell), we obtain a steeper variation of the LDL-like fraction as a function of the temperature, similarly to what we obtained with our model. This confirms that the discrepancy between the fraction estimated by the model and by the order parameters arises from the dimension of the minimal nucleus of molecules that defines either the LDL- or the HDL-like phase. Indeed, N_c and the coarse-graining radius N regulate the steepness of the crossover between the two structural regimes in a similar way: a larger size for the minimal cluster leads to a sharper crossover. It can be observed in the insets of panels (b) and (d) of Fig. 4 that coarse-graining on an increasing number of nearest neighbors has a very relevant effect on the LDL-fraction up to $N = 10$–12. A further enlarging of the spatial coarse-graining has instead a minor effect. The significant variation of the LDL-like fraction for $N \leqslant N_c \approx 15$ suggests that the N_c value obtained within our model is a reasonable value for the minimal cluster necessary to define the thermodynamic properties of the two liquid phases. In addition, the above-mentioned results also suggest an interesting connection between the N_c parameter of our thermodynamic model and structural order parameters that will be further investigated in future works.

IV. CONCLUSIONS

In this work, we apply a theoretical model we have recently proposed for fluid state systems[21] to the investigation of simulated supercooled water. The model uses basic statistical mechanical and thermodynamic relations to obtain the thermodynamic properties of liquid water as a function of temperature. The model shows that in supercooled conditions, water thermodynamics along isobars (i.e., as a function of the temperature) can be well described using two gamma state solutions (i.e., solutions based on Gamma distributions) representing the two liquid states differing in density: the LDL-like state and the HDL-like state. Remarkably, the very same description used to model the experimental data on the liquid- and gas-like thermodynamics also applies for simulated data in the supercooled region, confirming the strong parallelism between the pseudo-phase transition in real supercritical water and that in simulated supercooled water.

When compared to other two state models, the model we use here has more or less the same number of fitting parameters. However, while in other models some of these parameters come from a phenomenological/empirical description, in our model the fitting function strictly derives from a physical model. In addition, the linearity of the enthalpy when the system is in a single Gamma state furnishes a simple and efficient diagnostic criterion to identify the temperature range in which a single phase exists and to evaluate the accuracy of the model itself.

By interpreting the liquid system as an ensemble of interconverting LDL-like and HDL-like subsystems containing N_c molecules, the model also describes the transition between the LDL-like and the HDL-like states by crossing the Widom line. Application of the model using several temperature/pressure conditions in the supercooled region provides a series of Widom line temperatures in good agreement with those estimated by previous models.[24,25]

The fit of the enthalpy at the three pressures crossing the Widom line provides values for N_c, the minimum number of water molecules necessary to define a cluster that can be found in either the LDL- or the HDL-like state, of the same order of magnitude (≈10–20 molecules). Considering that the parameter N_c is completely unbiased, this suggests that in liquid water, the first hydration shell is not sufficient to distinguish between the two regimes in terms of statistical fluctuations. We also compare the LDL-like fraction provided by the model with that provided by structural order parameters, showing that this quantity strongly depends on the inclusion of medium-range order effects (up to the second shell). As a matter of fact, the LDL fraction estimated by the order parameters is in good agreement with that estimated by the model only if a coarse graining of the order parameter is performed beyond the first hydration shell.

SUPPLEMENTARY MATERIAL

Details on the MD simulations, the model fitting parameters, the enthalpy trends fit with a fixed N_c value, the radial distributions functions, the c_p and k_T trends obtained from the MD data, the distributions of ψ and ζ, and a discussion of the hydrogen bond identifications are included in the supplementary material.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Isabella Daidone: Conceptualization (equal); Investigation (equal); Writing – review & editing (equal). **Riccardo Foffi**: Conceptualization (equal); Formal analysis (equal); Writing – review & editing (equal). **Andrea Amadei**: Conceptualization (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Laura Zanetti-Polzi**: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] P. Gallo, K. Amann-Winkel, C. A. Angell, M. A. Anisimov, F. Caupin, C. Chakravarty, E. Lascaris, T. Loerting, A. Z. Panagiotopoulos, J. Russo *et al.*, "Water: A tale of two liquids," Chem. Rev. **116**, 7463–7500 (2016).
[2] P. H. Poole, F. Sciortino, U. Essmann, and H. E. Stanley, "Phase behaviour of metastable water," Nature **360**, 324–328 (1992).
[3] F. Sciortino, I. Saika-Voivod, and P. H. Poole, "Study of the ST2 model of water close to the liquid–liquid critical point," Phys. Chem. Chem. Phys. **13**, 19759–19764 (2011).
[4] Y. Li, J. Li, and F. Wang, "Liquid–liquid transition in supercooled water suggested by microsecond simulations," Proc. Natl. Acad. Sci. U. S. A. **110**, 12209–12212 (2013).
[5] Y. Ni and J. Skinner, "Evidence for a liquid-liquid critical point in supercooled water within the E3B3 model and a possible interpretation of the kink in the homogeneous nucleation line," J. Chem. Phys. **144**, 214501 (2016).
[6] J. C. Palmer, P. H. Poole, F. Sciortino, and P. G. Debenedetti, "Advances in computational studies of the liquid–liquid transition in water and water-like models," Chem. Rev. **118**, 9129–9151 (2018).
[7] T. E. Gartner III, K. M. Hunter, E. Lambros, A. Caruso, M. Riera, G. R. Medders, A. Z. Panagiotopoulos, P. G. Debenedetti, and F. Paesani, "Anomalies and local structure of liquid water from boiling to the supercooled regime as predicted by the many-body MB-pol model," J. Phys. Chem. Lett. **13**, 3652–3658 (2022).
[8] V. Holten, J. C. Palmer, P. H. Poole, P. G. Debenedetti, and M. A. Anisimov, "Two-state thermodynamics of the ST2 model for supercooled water," J. Chem. Phys. **140**, 104502 (2014).
[9] V. Holten and M. Anisimov, "Entropy-driven liquid–liquid separation in supercooled water," Sci. Rep. **2**, 713 (2012).
[10] F. Caupin and M. A. Anisimov, "Minimal microscopic model for liquid polyamorphism and waterlike anomalies," Phys. Rev. Lett. **127**, 185701 (2021).
[11] H. Tanaka, "Simple physical model of liquid water," J. Chem. Phys. **112**, 799–809 (2000).
[12] M. J. Cuthbertson and P. H. Poole, "Mixturelike behavior near a liquid-liquid phase transition in simulations of supercooled water," Phys. Rev. Lett. **106**, 115706 (2011).
[13] K. Wikfeldt, A. Nilsson, and L. G. Pettersson, "Spatially inhomogeneous bimodal inherent structure of simulated liquid water," Phys. Chem. Chem. Phys. **13**, 19918–19924 (2011).
[14] J. Russo and H. Tanaka, "Understanding water's anomalies with locally favoured structures," Nat. Commun. **5**, 3556 (2014).

[15]J. M. Montes de Oca, F. Sciortino, and G. A. Appignanesi, "A structural indicator for water built upon potential energy considerations," J. Chem. Phys. **152**, 244503 (2020).
[16]A. Neophytou, D. Chakrabarti, and F. Sciortino, "Topological nature of the liquid–liquid phase transition in tetrahedral liquids," Nat. Phys. **18**, 1248–1253 (2022).
[17]C. Faccio, M. Benzi, L. Zanetti-Polzi, and I. Daidone, "Low- and high-density forms of liquid water revealed by a new medium-range order descriptor," J. Mol. Liq. **355**, 118922 (2022).
[18]R. Foffi and F. Sciortino, "Correlated fluctuations of structural indicators close to the liquid–liquid transition in supercooled water," J. Phys. Chem. B **127**, 378 (2022).
[19]A. Mondal, G. Ramesh, and R. S. Singh, "Manifestations of the structural origin of supercooled water's anomalies in the heterogeneous relaxation on the potential energy landscape," J. Chem. Phys. **157**, 184503 (2022).
[20]A. Offei-Danso, A. Hassanali, and A. Rodriguez, "High-dimensional fluctuations in liquid water: Combining chemical intuition with unsupervised learning," J. Chem. Theory Comput. **18**, 3136–3150 (2022).
[21]L. Zanetti-Polzi, I. Daidone, and A. Amadei, "A general statistical mechanical model for fluid system thermodynamics: Application to sub-and super-critical water," J. Chem. Phys. **156**, 044506 (2022).
[22]A. Amadei, M. Apol, A. Di Nola, and H. Berendsen, "The quasi-Gaussian entropy theory: Free energy calculations based on the potential energy distribution function," J. Chem. Phys. **104**, 1560–1574 (1996).
[23]A. Amadei, M. Apol, and H. Berendsen, "Extensions of the quasi-Gaussian entropy theory," J. Chem. Phys. **106**, 1893–1912 (1997).
[24]R. S. Singh, J. W. Biddle, P. G. Debenedetti, and M. A. Anisimov, "Two-state thermodynamics and the possibility of a liquid-liquid phase transition in supercooled TIP4P/2005 water," J. Chem. Phys. **144**, 144504 (2016).
[25]J. W. Biddle, R. S. Singh, E. M. Sparano, F. Ricci, M. A. González, C. Valeriani, J. L. Abascal, P. G. Debenedetti, M. A. Anisimov, and F. Caupin, "Two-structure thermodynamics for the TIP4P/2005 model of water covering supercooled and deeply stretched regions," J. Chem. Phys. **146**, 034502 (2017).
[26]P. Debenedetti, F. Sciortino, and G. Zerze, "Second critical point in two realistic models of water," Science **369**, 289–292 (2020).
[27]J. Patel, C. Kapadia, and D. Owen, *Handbook of Statistical Distributions* (M. Dekker, 1976).
[28]V. Holten, D. T. Limmer, V. Molinero, and M. A. Anisimov, "Nature of the anomalies in the supercooled liquid state of the mW model of water," J. Chem. Phys. **138**, 174501 (2013).
[29]P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, 1953).
[30]J. L. Abascal and C. Vega, "A general purpose model for the condensed phases of water: TIP4P/2005," J. Chem. Phys. **123**, 234505 (2005).
[31]E. Shiratani and M. Sasai, "Growth and collapse of structural patterns in the hydrogen bond network in liquid water," J. Chem. Phys. **104**, 7671–7680 (1996).
[32]J. R. Errington and P. G. Debenedetti, "Relationship between structural order and the anomalies of liquid water," Nature **409**, 318–321 (2001).
[33]R. Foffi and F. Sciortino, "Structure of high-pressure supercooled and glassy water," Phys. Rev. Lett. **127**, 175502 (2021).
[34]F. Martelli, "Unravelling the contribution of local structures to the anomalies of water: The synergistic action of several factors," J. Chem. Phys. **150**, 094506 (2019).
[35]F. Martelli, H.-Y. Ko, E. C. Oğuz, and R. Car, "Local-order metric for condensed-phase environments," Phys. Rev. B **97**, 064105 (2018).
[36]A. K. Soper and M. A. Ricci, "Structures of high-density and low-density water," Phys. Rev. Lett. **84**, 2881 (2000).
[37]S. Fanetti, A. Lapini, M. Pagliai, M. Citroni, M. Di Donato, S. Scandolo, R. Righini, and R. Bini, "Structure and dynamics of low-density and high-density liquid water at high pressure," J. Phys. Chem. Lett. **5**, 235–240 (2014).
[38]R. Shi, J. Russo, and H. Tanaka, "Origin of the emergent fragile-to-strong transition in supercooled water," Proc. Natl. Acad. Sci. U. S. A. **115**, 9444–9449 (2018).