



# A journey into the local structural order of liquid water: from the insights earned by geometrically-inspired descriptors to the development of a brand new energy-based indicator

Alejandro R. Verde, Laureano M. Alarcón, Gustavo A. Appignanesi<sup>a</sup>

INQUISUR, Departamento de Química, Universidad Nacional del Sur (UNS)-CONICET, Avenida Alem 1253, 8000 Bahía Blanca, Argentina

Received: 14 August 2022 / Accepted: 22 September 2022

© The Author(s), under exclusive licence to Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2022

**Abstract** In this work, we review recent efforts to develop a structural order parameter that satisfactorily accounts for water's anomalous behavior. Over the years, a large body of evidences has been accumulated in favor of a two-liquids scenario for water. This picture considers the liquid as a mixture of two different local molecular arrangements (one well-structured, with low local density, and another unstructured, with high local density) and water's anomalies would stem precisely from the competition between such states. Thus, the availability of a structural indicator capable of accurately discriminating both kinds of molecular arrangements is mandatory within this field and has been the subject of intense work during the past decades. In turn, comparative reports of the methods for local structure determination in liquids and glasses, including liquids with directional interactions like water, have already been performed. Thus, instead of presenting a complete account of such methods for water, we shall now focus on structural indicators that have been explicitly devised to sense the existence of the two purported water states. In this sense, we shall present indicators that have exhibited success in producing distributions with certain degree of bimodality. Some of these indicators are based on the consideration of certain structural traits like the tetrahedral order of water's first coordination shell or the translational order up to the second shell (its inhomogeneity). However, we shall show that besides the inability of early indicators to produce clear bimodal distributions (which would enable not only the estimation of the population of the two kinds of species but that would also allow for a proper molecular classification), such indices have also exhibited a tendency to overestimate the unstructured state. We shall thus also describe strategies to overcome these limitations, like the employment of potential energy minimizations and the incorporation of water's intermolecular interactions (hydrogen bonds), which have either improved existent indicators or proposed new ones. Within this context, a very recent proposal entirely based on energetic considerations will also be described.

## 1 Introduction

Water is a fascinating system with an enormous intrinsic relevance and that is also central for many main scientific and technological fields [1–25]. In particular, and at variance from simple liquids, water has been shown to present a plethora of thermodynamic and dynamic anomalies which make it irreplaceable for the structure and dynamics of systems that range from biology to materials science [2, 3, 26, 27]. This behavior has been proposed to stem from the existence of a second critical point located within the liquid supercooled regime [3] (the supercooled liquid represents a metastable regime achieved when a liquid is cooled fast enough to prevent crystallization so that its dynamics slows down and its viscosity increases drastically without displaying significant structural changes [1, 28, 29]). This two-liquids picture [2, 3, 26, 27] was inspired by water's polyamorphism in the glass regime: the existence of a low-density amorphous ice (LDA) and a high-density amorphous (HDA) ice [2, 3, 5]. Under this scenario, the HDA–LDA coexistence line might continue in the liquid regime until ending at a liquid-liquid critical point [2, 3]. The structures of these liquid states, one with low and the other with high local density environments, might thus resemble, respectively, that of LDA and HDA, and the anomalous behavior would result from the competition of these two stable local molecular arrangements. We can also speak of the LDL and HDL states as structured and unstructured, respectively, since the LDL should be represented by correctly tetrahedrally-coordinated molecules, while the HDL would imply the presence of molecules whose tetrahedral structuring is partially distorted to the point that one of its hydrogen bonds might be (at least partially) broken. In well-structured molecules there should exist a clear gap between the first and second coordination shell, while for the unstructured ones, the intrusion of interstitial molecules from the second shell would perturb the local ordering of the first one. A fifth neighbor located at an interstitial position could also be involved in the formation of a bifurcated hydrogen bond, a motif that might promote mobility, since the rearrangement of the local hydrogen-bond network would be eased provided that one hydrogen bond could be forming while the one replaced is still elongating

<sup>a</sup> e-mail: [appignan@criba.edu.ar](mailto:appignan@criba.edu.ar) (corresponding author)

(thus removing the need for a net bond rupture with its concomitant energy barrier) [30]. While abundant computational [31–34] and experimental [35–40] evidences for the existence of such two liquid phases has been produced, the unambiguous determination of their precise nature and the quantification of their relative abundances has been difficult, particularly at liquid temperatures where thermal fluctuations promote deformations of the local structures. Experiments face the problem of the intervention of homogeneous crystal nucleation and, thus, molecular dynamics simulations have been central in this field. Within this context, it becomes mandatory to possess structural indicators to quantify the two water molecular states and, if possibly, to classify the water molecules. Several indices have been proposed during the years and applied in molecular dynamics simulation studies [8, 22, 41–54] (for an excellent review on methods for local structure determination in liquids and glasses, including liquids with directional interactions like water, please refer to [53]). Generally, these indicators are built upon certain geometrical characteristics of the water molecules, like the degree of tetrahedral order of the first shell or the translational order up to the second shell. However, it has been difficult to find a structural order parameter that directly correlates with water's anomalies [8]. Additionally, most efforts have not been able to produce clear bimodal distributions that, beyond the estimation of the fraction of the two kinds of molecules, would enable a proper classification of each of the water molecules in the sample. Moreover, as we shall see further on, structurally-inspired methods tend to overestimate the high-density state. Thus, instead of intending a complete review on methods for structural determination (since, as indicated, excellent reviews are already at hand [53]), we shall now focus mainly on the strategies taken to produce bimodality (while rationalizing the failures of certain indices to produce it), like the introduction of potential energy minimizations and the explicit consideration of hydrogen-bonding. Thus, we shall make special emphasis on the indicators that have indeed succeeded in producing such bimodality. Among them, we shall consider a recently introduced brand new structural indicator that, abandoning geometrical preconceptions, is explicitly based on energetic considerations. This indicator has been shown to accurately discriminate and classify water molecules, to produce radial distribution functions resembling their HDA and LDA analogs and to nicely explain the anomalous behavior of water's isobaric heat capacity.

## 2 Early indicators and the relevance of introducing potential energy minimizations

Some early indicators aimed at revealing local structural order in water were based on the particular arrangement of the molecular first coordination shell as produced by the hydrogen bonds. Thus, they aimed at measuring the degree of local tetrahedral order of the water molecules, as done for example by the parameters  $q$  [41, 42] and  $\theta$  [43, 44]. The tetrahedral index  $q$ , is defined as follows:

$$q(i, t) = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 \left( \frac{1}{3} + \cos \psi_{jk}(i, t) \right)^2,$$

where  $\psi_{jk}(i, t)$  is the angle between the lines connecting the oxygen of molecule  $i$  with those of its nearest molecules  $j$  and  $k$  at time  $t$ . Then,  $-3 \leq q(i, t) \leq 1$ , and a high value indicates that molecule  $i$  has a good tetrahedral order [41, 42].

The tetrahedrity index  $\theta$ , in turn, has the following mathematical form:

$$\theta(i, t) = \frac{6 \sum_{j=1}^5 \sum_{k=j+1}^6 [l_j(t) - l_k(t)]^2}{15 \sum_{j=1}^6 l_j^2(t)},$$

where  $l_k(t)$  are the lengths of the six ( $k \in \mathbb{N}$ ,  $1 \leq k \leq 6$ ) edges of the tetrahedron formed by the four nearest (oxygens) neighbors from oxygen  $i$  (oxygen  $i$  is located in the center of the tetrahedron). For an ideal tetrahedron,  $\theta(i, t)$  is equal to zero.

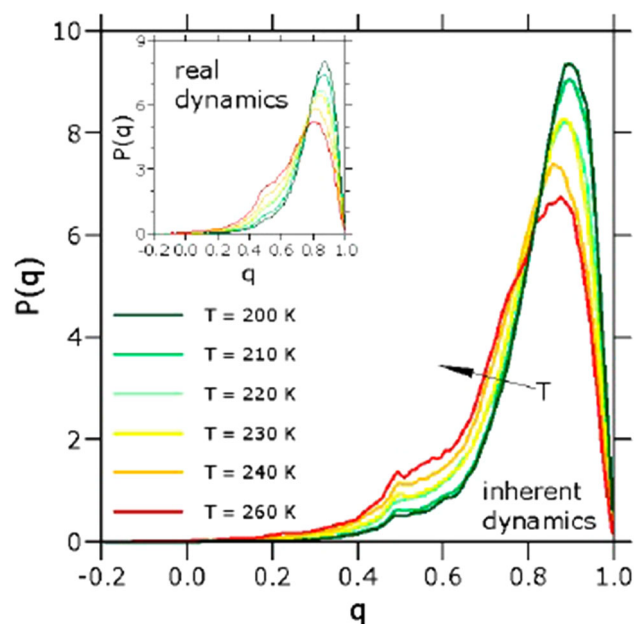
Others indicators, like the Local Structure Index, LSI [45, 46] or the index  $d_5$  [51], measured the translational order up to the second coordination shell. This fact was motivated by the observation of the existence of a clear gap between the first and second coordination shells (between around 3.2 Å and 3.8 Å in the radial neighbor distribution) for well-structured molecules while the unstructured ones exhibit the presence of interstitial molecules (as for example a fifth neighbor) that might disturb the order of the first shell. The parameter  $d_5$  is calculated simply as the distance from the oxygen atom of a central molecule and its fifth nearest neighbor [51]. The LSI, in turn, is built in the following way [45, 46]: For each molecule  $i$  one orders the rest of the molecules depending on the radial distance  $r_j$  between the oxygen of the molecule  $i$  and the oxygen of molecule  $j$ :  $r_1 < r_2 < r_j < r_{j+1} < \dots < r_{n(i,t)} < r_{n(i,t)+1}$ ,  $n(i)$  is chosen so that  $r_{n(i)} < 3.7 \text{ \AA} < r_{n(i)+1}$ . Then,  $\text{LSI}(i) = I(i)$  is defined as [45, 46]:

$$I(i) = \frac{1}{n(i, t)} \sum_{j=1}^{n(i)} [\Delta(j; i, t) - \bar{\Delta}(i)]^2$$

where  $\Delta(j; i) = r_{j+1} - r_j$  and  $\bar{\Delta}(i)$  is the average over all molecules of  $\Delta(j; i)$ . Thus,  $I(i)$  expresses the inhomogeneity in the radial distribution within the sphere of radius around 3.7 Å. A high value of  $I(i)$  implies that molecule  $i$  at the present configuration is characterized by a tetrahedral local order and a low-local density, while on the contrary, values of  $I(i) \approx 0$  indicate a molecule with defective tetrahedral order and high-local density.

All these indicators provided first evidence for the picture of water as a mixture of two different kinds of local molecular arrangements, since their distributions indeed exhibited certain mild bimodal features. Deconvolution procedures could then be

**Fig. 1** Probability density  $P(q)$  of finding a molecule with tetrahedral index  $q$  for the SPC/E model at temperatures 200 K, 210 K, 220 K, 230 K, 240 K and 260 K, evaluated at the inherent structures (minimized configurations). Inset: The same curves but using the real dynamics. With kind permission of The European Physical Journal (EPJ) [47]

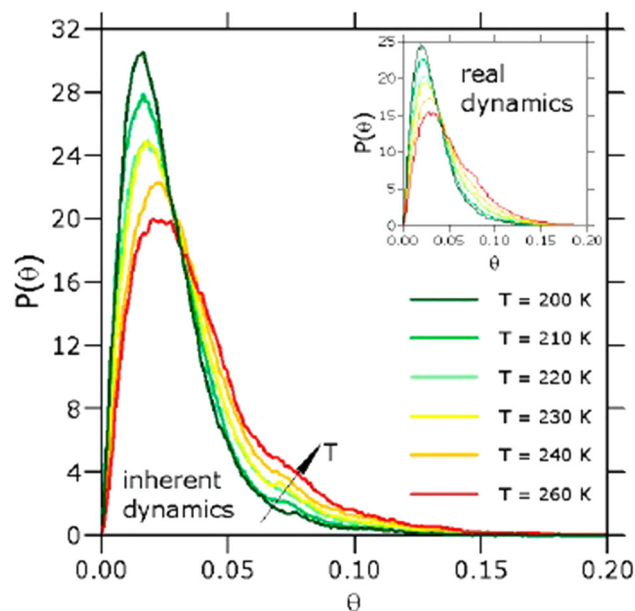


employed to estimate the populations of the two states. However, none of them, nor the total energy of the molecule [55] which was also employed as an structural indicator, produced clear bimodal distributions (with a deep minimum separating the two expected states) for the different water models studied [47, 48, 52]. Besides enabling a more accurate estimation of the two populations, clear bimodal distributions would also allow for molecular classification. Another problem presented by these indices is the overestimation of the population of the unstructured state. This is very notorious, for example, for the LSI index. This fact might partially be due to the masking effect of thermal fluctuations that prevent a proper identification of ordered local structures. If the temperature is low in the supercooled regime, it is expected that the local molecular arrangements are not significantly deformed. However, at higher temperatures (more importantly close or above the melting temperature), thermal fluctuations can induce geometrical deformations that blur the traits involved in the structural determination. Distances and angles can be significantly modified and even hydrogen-bond coordination defects can be confused with locally deformed arrangements. Thus, a relevant advance in this field was the decision to study them under the inherent structures approach, that is, to resort to potential energy minimizations. This method [56–59] applies a minimization process to partition the potential energy surface (PES) of the system into disjoint basins. Such basins represent the set of configuration space points that belong to the same local minimum which is called inherent structure (IS) when the configurations are subject to a minimization process. Thus, starting from a given instantaneous structure of the molecular dynamics trajectory (at the real dynamics) one minimizes potential energy (for example, by means of the steepest-descent algorithm) until convergence, in order to attain the corresponding local basin of attraction in the PES [56–59]. However, it has been shown that the application of this procedure does not produce significant changes for some of the geometrically-based indicators above-expounded [47, 48, 52]. For example, Figs. 1 and 2 display, respectively, the probability distributions of the indices  $q$  and  $\theta$  for the Simple Point Charge (Extended) potential (SPC/E) water model [60]. In both cases, the parameters are calculated at the real dynamics (instantaneous configurations) and also at the inherent dynamics (after minimization, that is, after attaining the corresponding IS, filtering out the randomizing effect of the thermal fluctuations as already indicated).

However, the situation was completely different for the case of the LSI indicator [47, 48, 61, 62]. Figure 3 shows the distribution of the LSI index for one temperature above and one temperature below the melting temperature of the corresponding water model (for estimation of melting temperatures please see [63]). Clear bimodal distributions with a deep minimum separating the two maxima are obtained for the inherent structures. In turn, no neat bimodality is evident at the real dynamics. For the IS, the peak at the left side (with an amplitude that increases with temperature) corresponds to unstructured (HDL-like) water molecules and the peak on the right corresponds to highly structured (LDL-like) molecules. Notably, the position of the minimum between the two peaks has been shown not to depend on temperature (a kind of “isosbestic point”), thus indicating the existence of two distinct populations in dynamical equilibrium [47]. Hence, besides being able to quantitatively estimate the fraction of both kinds of structures, the method provides us with a central evident advantage: It enables the classification of each of the water molecules. This made the LSI at the inherent dynamics scheme to be widely used both within the supercooled regime and for the liquid above the melting point and the same kind of bimodality was observed for all the other water models where it has been applied so far, like TIP4P/2005 [64, 67], TIP4P-ice [65, 66], TIP5P-Ew [50, 68] and also for ab initio potentials within the density functional framework [69].

The success of the strategy of combining the LSI indicator with potential energy minimizations can be rationalized to arise from the fact that the index directly senses the existence or not of a gap between the first and second molecular coordination shells [61, 62]. The minimization procedure cleans the inter-shell region since it improves the arrangement of both shells. This would be

**Fig. 2** Probability density distribution of the tetrahedrity index  $\theta$ ,  $P(\theta)$ , for the same case as that of Fig. 1. With kind permission of The European Physical Journal (EPJ) [47]



the case both for well-structured molecules and for situations when the inter-shell occupancy at the real dynamics is due to mere thermal deformations [61, 62]. However, for a true interstitial molecule where there is certain energetic gain (local minimum) in residing in the inter-shell region (thus implying a genuine structural distortion), the minimization would not retrieve it to its original coordination shell [61, 62].

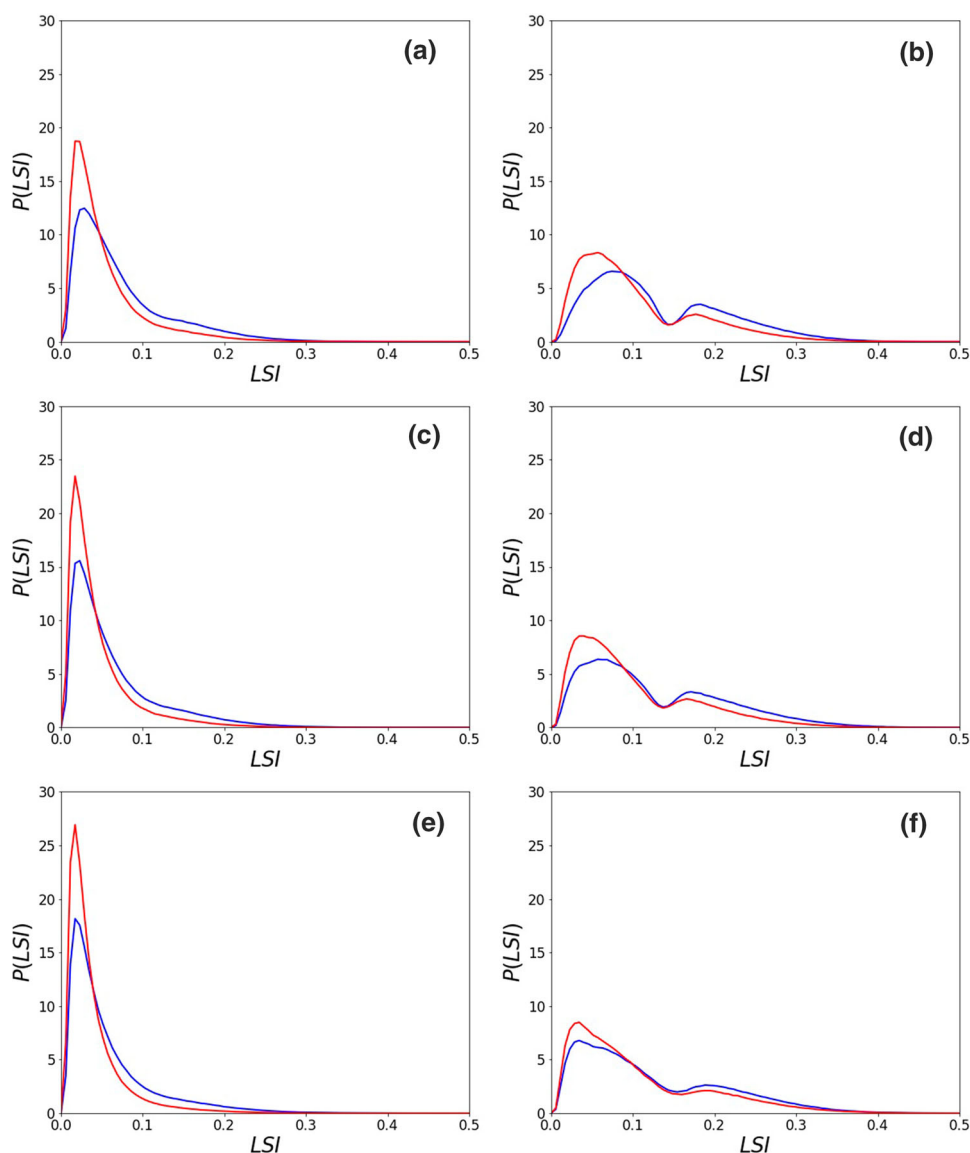
Despite the wide popularity gained by this method, a couple of caveats remained. In one place, the position of the minimum and the relative populations of both kinds of species depend on the arbitrary threshold used to define the index (originally 3.7 Å) [48]. On the other hand, the population of the unstructured state is found to be dominant even in a great part of the supercooled regime (below the melting point), a fact that might speak of an underestimation of the structured state. This situation also holds at different extent for all the rest of the indicators above-mentioned (and of course it is worse for the LSI in the case it is calculated at the real dynamics). But even when for the LSI index this failure is partially corrected by the minimization procedure, its performance is still suboptimal within such context.

### 3 The explicit inclusion of hydrogen bonding

More recently, a new indicator that considers the role of water's intermolecular interactions was introduced,  $\zeta$  index [8, 52]. While still based on the quantification of the translational order up to the second coordination shell, like the LSI index and  $d_5$ , this new indicator explicitly incorporates hydrogen bonding albeit still in a geometric manner: it measures the distance between the last molecule hydrogen-bonded to the central one (the farthest hydrogen-bonded neighbor, called LHB) and the closest molecule not hydrogen-bonded to the central one (the first non-hydrogen-bonded one, 1NHB). In most cases (for example for correctly four-fold hydrogen-bond coordinated molecules), this would be the distance from the fourth to the fifth neighbors. This index, which implied a conceptual advance over the previous ones, produced fairly bimodal distributions for the TIP5P model of water directly at the instantaneous configurations (that is, at the real dynamics, without the need to resort to minimization techniques) [52] (see Fig. 4). However, the bimodality was less neat for TIP4P/2005 and even less for SPC/E [8, 52, 61, 62]. Moreover, the employment of a minimization procedure has been shown not to improve this situation [62], as can be learnt from Fig. 4.

Within this context, it is instructive to learn on the effect of the minimization procedure on the two distances involved in  $\zeta$  index calculation. Figure 5 shows that the minimization is indeed vacating the region in between the two first two coordination shells [62]. In certain cases, the fifth neighbor is displaced outwards while the order of the first coordination shell is improved. In other cases, however, instead of considering the fourth and fifth neighbors as usually expected, the index is in fact dealing with the third and the fourth ones since the central molecule is three-fold hydrogen-bond coordinated at the real dynamics. In some of these instances, the fourth neighbor is displaced away from the central molecule by the minimization procedure, thus implying that there exists a genuine defect in the hydrogen-bond coordination: the central molecule is a truly three-fold coordinated molecule, or defect. However, in other cases the fourth molecule is moved back to distances within the first coordination shell, thus implying that the fourth hydrogen bond is restored. These molecules, highly deformed by thermal fluctuations, would be classified as HDL-like on geometrical terms while the minimization procedure is telling us that they indeed belong to an LDL-like basin of attraction. This fact represents a warning against geometrically-based indicators since the structured state could be clearly underestimated (it is not surprising, thus,

**Fig. 3** Probability density  $P(I)$  of finding a molecule with local structure index  $LSI = I$  for two temperatures, above and below the melting point for SPC/E water model ( $T = 240$  K, red curve, and  $T = 210$  K, blue curve, respectively). The case **a** corresponds to the real dynamics while **b** is for the inherent dynamics. At the inherent dynamics, neat bimodal distributions are evident. The curves cross at a value  $I_{\min} \approx 1.4 \cdot 10^{-3} \text{ nm}^2$ . When the same configurations are evaluated at the real dynamics, no clear separation in two populations is observable. We also show the cases for water models TIP4P/2005 [cases (c) and (d), at  $T = 246$  K, blue curve, and  $T = 281$  K, red curve] and TIP5P [cases (e) and (f), at  $T = 267$  K, blue curve, and  $T = 305$  K, red curve]. The curves on the left (a, c, e) correspond to the real dynamics while the ones on the right are for the inherent dynamics. As indicated, for each case, the red curve corresponds to the higher temperature and the blue one is for the lower temperature

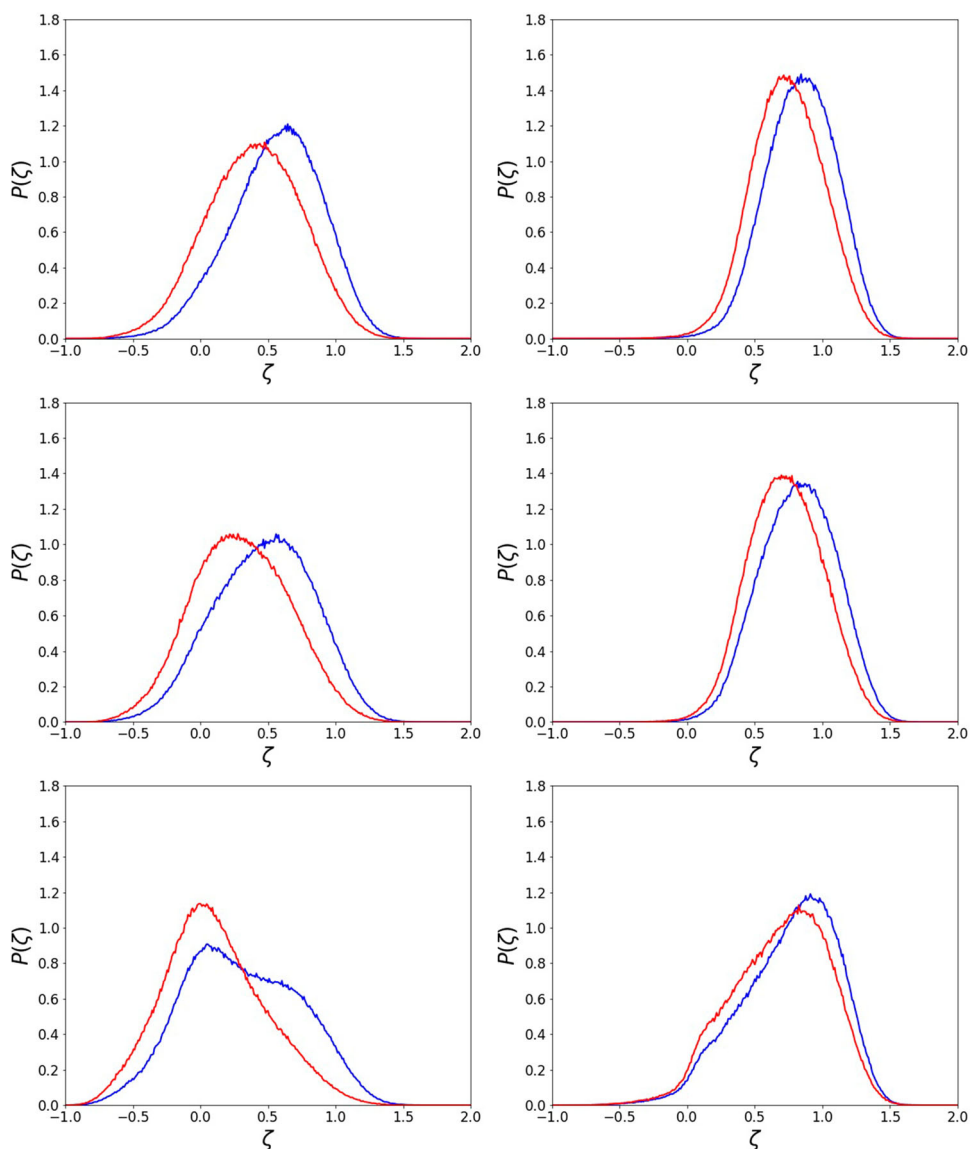


that  $\zeta$  index does not exhibit a clear dominance of the population of the structured state at the beginning of the supercooled regime, that is, at temperatures when the melting point is crossed). From the considerations made, it is evident that  $\zeta$  index represented a clear conceptual advance which explicitly introduced a key ingredient which is expected to lay at the heart of water's anomalies (the intermolecular interactions). However, this incorporation of the role of hydrogen bonding was still performed in a geometric way. Hence, the knowledge gained from the performance of this indicator in different situations and its response to potential energy minimizations, provided inspiration for the development of a brand new energy-based indicator.

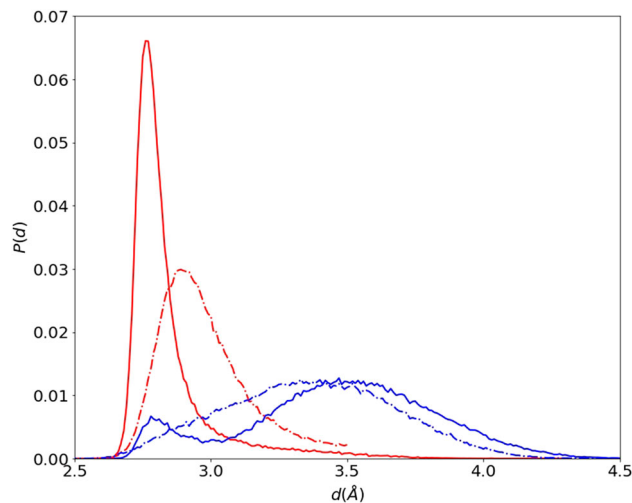
#### 4 A novel energy-based structural indicator

As above expounded, the molecular-level information gained from the application of  $\zeta$  index to different water models and from its response when subject to minimization, highlighted the need to abandon structural preconceptions and to devise an indicator based exclusively on the consideration of water's molecular interactions. This fact, thus, inspired the development of  $V_4$  indicator [54]. To determine the value of this index for a water molecule  $i$  one calculates all its pair-wise interactions  $V_{ij}$ ,  $j \neq i$ , which are then sorted in increasing order regarding their intensity.  $V_4(i)$  is thus defined as the fourth  $V_{ij}$ . For a molecule with four well-formed hydrogen bonds  $V_4(i)$  would be close to the linear hydrogen bond energy (large negative value) while for a molecule which presents interstitial neighbors disturbing at least one of its four hydrogen bonds  $V_4(i)$  would yield a larger (less negative) value. When applied

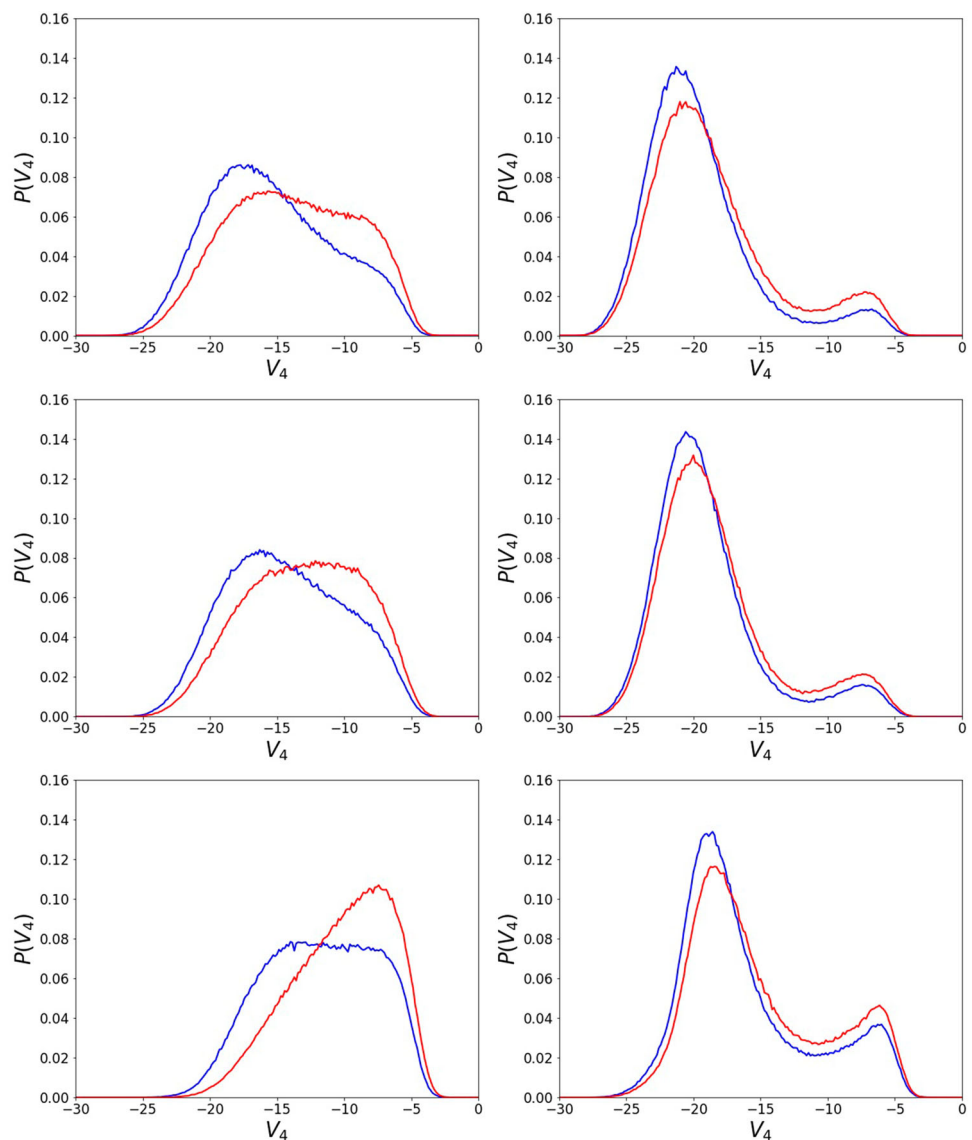
**Fig. 4** Probability distribution of  $\zeta$  index for different water models both at the real and inherent dynamics (curves at the left or at the right, respectively).  $\zeta$  index is given in Å. The curves are analogous to the ones of Fig. 3, that is, same systems and same temperatures (from top to bottom: SPC/E, TIP4P/2005 and TIP5P and, as always, the red curves correspond to the higher temperature while the blue ones are for the lower temperature)



**Fig. 5** Distance from the central molecule to the 1NHB and LHB molecules for SPC/E at  $T = 240$  K when the molecules are classified in the real dynamics and also the distances are measured in such instantaneous configurations (1NHB blue dotted curve and LHB red dotted curve) and when the molecules are classified in the real dynamics but their distances to the central molecule are calculated in the corresponding inherent structures, that is, after minimizing the instantaneous configurations (1NHB, blue solid curve, and LHB, red solid curve)



**Fig. 6** Probability distribution for  $V_4(i)$  for a water molecule for a couple of temperatures, one above and the other below the melting point (for the water model SPC/E at a pressure of 1 bar).  $V_4$  is given in units of KJ/mol. We also depict the situation for the real trajectory (without minimization). The curves are analogous to the ones of Figs. 3 and 4, that is, same systems and same temperatures (from top to bottom: SPC/E, TIP4P/2005 and TIP5P; as always: curves at the left are for the real dynamics and the ones at the right are for the inherent structures and the red curves correspond to the higher temperature while the blue ones are for the lower temperature)

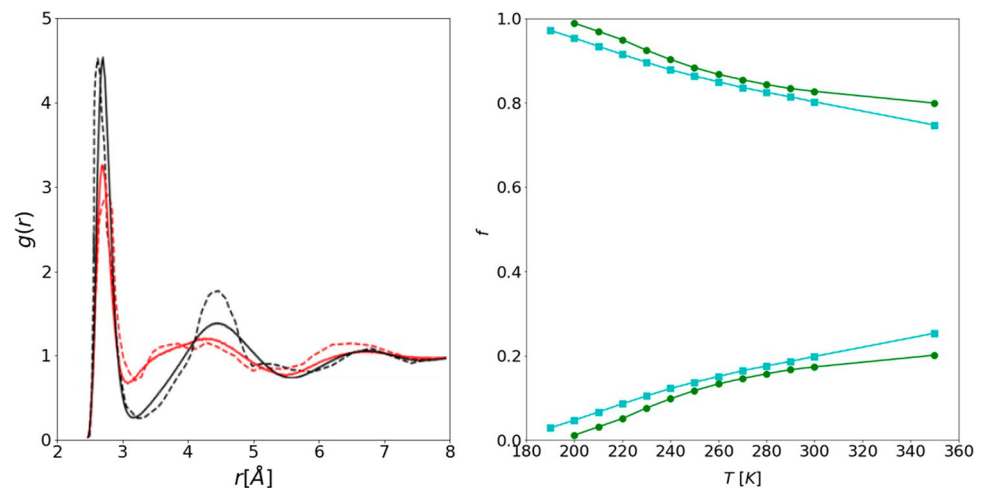


to different water models, like SPC/E, TIP4P/2005 and TIP5P,  $V_4$  exhibited neat bimodal distributions [54, 70] for the liquid regime both above the melting point and within the supercooled regime.

In Fig. 6 we show results for different water models in the real dynamics and for the inherent structures. A neat bimodality is evident in both schemes, but preferentially at the inherent dynamics. The left-side peak comprises molecules with four strong tetrahedrally-arranged interactions. These molecules were thus called  $T$  molecules. In turn, the peak on the left involves molecules where the first coordination shell is distorted, thus preventing the formation of a good quality fourth hydrogen bond. Hence, such molecules were regarded as undercoordinated molecules or defects ( $D$  molecules) [54]. The inherent dynamics scheme is the appropriate one for classification purposes since it is able to discriminate between genuinely distorted molecules (defects) and merely thermally deformed ones [54, 70]. As we can learn from direct inspection of Fig. 6, the  $T$  molecules represent the dominant state at both temperatures while the  $D$  ones indeed correspond to the minority component typical of defect configurations. Thus,  $V_4$  not only accurately estimates the fractions of the two competing local molecular arrangements but it properly classifies water molecules within these two kinds of species, a result that had been out of reach for previous indicators [54]. The ability of  $V_4$  to correct the tendency of previous indices to overestimate the fraction of unstructured molecules and the possibility to accurately detect them constitutes a clear breakthrough since such defects are paramount both for structure characterization (as for example in determining the salient features of the radial distribution function) and for dynamics (such defects promote the hydrogen bond rearrangement events responsible for water's relaxation dynamics [30, 54, 70–72]).

In Fig. 7, we show an example of radial distribution functions where a great resemblance between  $T$  and LDA molecules and also between  $D$  and HDA molecules is evident. Such a close similarity to the amorphous solids had not been obtained by previous indicators. We also depict the temperature-dependence of the population of the two kinds of water molecules. At variance from

**Fig. 7** Left: Oxygen–oxygen radial distribution functions (RDF) separated into two components ( $T$  or  $D$ ) according to the molecule located at the origin. Results are for the SPC/E model for  $T = 210$  K and 1 bar at the real dynamics scheme (we recall that the classification in  $T$  or  $D$  is always performed at the inherent dynamics in order to remove the masking effect of the thermal fluctuations). We compare with the experimental HDA and LDA results reported (at 80 K) in ref. [73]. Right: Fraction of  $T$  molecules (circles, majority component),  $f_T$ , and  $D$  molecules (squares, minority component),  $f_D$ , as a function of temperature. Green curves depict the case of TIP4P/2005 while light blue curves belong to the SPC/E model



other indicators that underestimate the structured LDA-like state, we can learn that the  $T$  molecules are dominant at all temperatures studied, thus again indicating the fact that water can be described as a random tetrahedral network with a few defects (D molecules) whose population is temperature-dependent. In this sense, the low-density liquid, LDL, would be a virtually pure  $T$  phase while the high density liquid, HDL, would consist of a mixture of a few D molecules surrounded by  $T$  ones, thus highlighting the need for the two-liquids scenario to go beyond the single-molecule description toward one based on multi-molecule local arrangements [54].

Additionally, the introduction of  $V_4$  has been central to shed light on some of water's thermodynamic and dynamic anomalies. On the one side, correlations of low- and high-dynamic-propensity molecules with  $T$  and  $D$  molecules, respectively, have been determined, thus providing evidence for the existence of the rather elusive structure-dynamics link in water's glassy relaxation [70]. On the other hand, the anomalous behavior of water's isobaric heat capacity ( $C_p$ ) at low temperature has been rationalized: it stems from the interconversion between  $T$  and  $D$  molecules with temperature, while each component displays a normal liquid behavior [54].

We note that in this review we have been mainly concerned with rigid water models. The study of flexible, polarizable and other more realistic water models would indeed be interesting. There are few studies that already considered some of the structural indicators for more realistic water models. In one case [74], the LSI distribution for the BK3 model is compared to that of the non-polarizable rigid TIP4P/Ice model, finding similar results (working at the real dynamics). More interestingly, another work [69] calculates the LSI at the inherent dynamics and shows similar results between an ab initio water model and also the TIP4P/Ice model.

Finally, it is also worth mentioning the recent increasing application of Machine Learning algorithms to systems with solid polymorphism, including water [75–81]. In particular, a couple of Neural Network approaches based in structural information have been applied to successfully classify water molecules between HDA- and LDA-like [75, 76]. The paper by Martelli et al. [76] is indeed very closely related to the results we show in Fig. 7. Interestingly, for TIP4P/2005 at a pressure close to 1 bar, they find a fraction of HDA-like molecules that goes from roughly 1% at  $T = 205$  K to around 20% at 250 K. These results are similar to the ones we get for the D molecules by employing  $V_4$  indicator (see Fig. 7), values much lower than the ones predicted by other structural indices that, as we have already indicated, tend to overestimate the fraction of HDA-like molecules.

## 5 Conclusions

The development of a structural order parameter that directly correlates with water's anomalies has not been easy, as recently recognized [8]. And in view of the two-liquids scenario [2, 3] that considers water as a mixture of two different local molecular arrangements or components (a structured low-local-density one, LDL, and a unstructured high-local-density one, HDL), the availability of an indicator that would produce clear bimodal distributions (capable of accurately classifying the water molecules besides estimating their relative populations), represents a matter of utmost relevance. In this work we have reviewed some of the recent efforts in such direction. Instead of intending a complete account of the structural descriptors developed along the years, we have focused mainly on structural indicators that have demonstrated success in providing certain features of bimodality in their distributions (there already exists a previous review on methods for determining structure in water and other liquids [53] and there are also other recent interesting propositions for water, like the local order metric (LOM), which uses the second neighbors sphere



[82, 83]). We have shown that after early efforts with mild success, a benchmark in this journey has been the application of the local structure index combined with potential energy minimizations (that is, at the inherent structure level). However, this index (as all previous ones based on geometrical considerations) underestimates the fraction of the structured component while it also depends on the definition of an arbitrary threshold value. We have also shown that a conceptual more recent step in this regard has been the introduction of another indicator also based on the consideration of the translational order up to the second shell,  $\zeta$  index, which explicitly included the consideration of water hydrogen bonds (the incorporation of hydrogen bonds would also be expected in terms of the purported role of such water intermolecular interactions in its anomalous behavior). However, the inclusion of these molecular interactions in  $\zeta$  index calculation was still performed at a geometrical level. Thus, the index only provided neat bimodal distributions for the TIP5P water model below the melting point, while this performance is less successful for other water models. We have also shown that a careful investigation of the local molecular arrangement that produce different values of  $\zeta$  index together with its response when subject to potential energy minimizations, provided clues to the problems faced by indicators built upon geometrical preconceptions (like the underestimation of the structured component) and thus fostered the need for the development of a new indicator based solely on energetic considerations. Hence, we have finally described the very recent introduction of a clear-cut parameter-free structural index,  $V_4$  indicator, that produced neat bimodal distributions for all water models studied. While additional work is still needed, this indicator has already provided us with valuable information for the two-liquids picture [54, 70]: It produced radial distribution functions of remarkable similarity with experimental low and high-density amorphous ices (pointing to the fact that the HDL–LDL equilibrium would be a thermally distorted version of the HDA–LDA one), it unravelled the underpinnings of water anomalous isobaric heat capacity and it provided evidence for the existence of the elusive link between structure and dynamics in the glassy regime.

**Acknowledgments** The authors acknowledge support from CONICET, UNS and ANPCyT (PICT2015/1893 and PICT2017/3127).

**Data Availability Statement** The data generated during the current study are available from the corresponding author on reasonable request.

## References

1. P.G. Debenedetti, *Metastable Liquids* (Princeton University Press, Princeton, 1996)
2. P.H. Poole, F. Sciortino, U. Essmann, H.E. Stanley, *Nature* **360**, 324 (1992)
3. P. Gallo et al., *Chem. Rev.* **116**, 7463–7500 (2016)
4. P.G. Debenedetti, H.E. Stanley, *Phys. Today* **56**, 40 (2003)
5. O. Mishima, H.E. Stanley, *Nature* **396**, 329 (1998)
6. C.A. Angell, *Chem. Rev.* **102**, 2627 (2002)
7. C.A. Angell, *Annu. Rev. Phys. Chem.* **55**, 559 (2004)
8. J. Russo, H. Tanaka, *Nat. Commun.* **5**, 3556 (2014)
9. P. Ball, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13327 (2017)
10. D.M. Huang, D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8324 (2000)
11. X. Huang, C.J. Margulis, B.J. Berne, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11953 (2003)
12. N. Giovambattista, P.G. Debenedetti, C.F. Lopez, P.J. Rossky, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2274 (2008)
13. D. Vitkup, D. Ringe, G.A. Petsko, M. Karplus, *Nat. Struct. Biol.* **7**, 34 (2000)
14. N. Choudhury, B. MontgomeryPettitt, *J. Phys. Chem. B* **109**, 6422 (2005)
15. G. Camisasca, A. Iorio, M. De Marzio, P. Gallo, *J. Mol. Liquids* **268**, 903 (2018)
16. M.H. Köhler, J.R. Bordin, L.B. da Silva, M.C. Barbosa, *Phys. A: Stat. Mech. Appl.* **490**, 331 (2018)
17. E.P. Schulz, L.M. Alarcón, G.A. Appignanesi, *Eur. Phys. J. E* **34**, 114 (2011)
18. S.R. Accordino, J.M.M. de Oca, J.A. Rodríguez Fris, G.A. Appignanesi, *J. Chem. Phys.* **143**, 154704 (2015)
19. C.A. Menéndez, S.R. Accordino, D.C. Gerbino, G.A. Appignanesi, *PLoS ONE* **11**, e0165767 (2016)
20. S.R. Accordino, D.C. Malaspina, J.A. Rodríguez Fris, L.M. Alarcón, G.A. Appignanesi, *Phys. Rev. E* **85**, 031503 (2012)
21. E. Schulz, M. Frechero, G. Appignanesi, A. Fernández, *PLoS ONE* **5**(9), e12844 (2010). <https://doi.org/10.1371/journal.pone.0012844>
22. S.R. Accordino, D.C. Malaspina, J.A. Rodríguez Fris, G.A. Appignanesi, *Phys. Rev. Lett.* **106**, 029801 (2011)
23. L.M. Alarcón, M. de los Angeles Frías, M.A. Morini, M.B. Sierra, G.A. Appignanesi, E. Anibal Disalvo, *Eur. Phys. J. E* **39**, 94 (2016)
24. S.R. Accordino, M.A. Morini, M.B. Sierra, J.A. Rodríguez Fris, G.A. Appignanesi, A. Fernández, *Proteins Struct. Bioinf.* **360**, 324 (2012)
25. J.M.M. de Oca, J.A.R. Fris, S.R. Accordino, D.C. Malaspina, G.A. Appignanesi, *Eur. Phys. J. E* **39**, 124 (2016)
26. P. Gallo et al., *Eur. Phys. J. E* **44**, 143 (2021). <https://doi.org/10.1140/epje/s10189-021-00139-1>
27. H.R. Corti et al., *Eur. Phys. J. E* **44**, 136 (2021). <https://doi.org/10.1140/epje/s10189-021-00136-4>
28. J.A. RodríguezFris, G.A. Appignanesi, E. La Nave, F. Sciortino, *Phys. Rev. E* **75**, 041501 (2007)
29. J.A. Rodríguez Fris, L.M. Alarcón, G.A. Appignanesi, *Phys. Rev. E* **76**, 011502 (2007). <https://doi.org/10.1103/PhysRevE.76.011502>
30. F. Sciortino, A. Geiger, H. Stanley, *Nature* **354**, 218 (1991)
31. J. Palmer, F. Martelli, Y. Liu et al., *Nature* **510**, 385–388 (2014)
32. P.G. Debenedetti, F. Sciortino, G.H. Zerze, *Science* **369**, 289–292 (2020)
33. F. Sciortino, I. Saika-Voivod, P.H. Poole, *Phys. Chem. Chem. Phys.* **13**, 19759 (2011)
34. J.C. Palmer, P.H. Poole, F. Sciortino, P.G. Debenedetti, *Chem. Rev.* **118**, 9129 (2018)
35. G. Walrafen, *J. Chem. Phys.* **47**, 114 (1967)
36. A. Nilsson, L.G. Pettersson, *Chem. Phys.* **389**, 1 (2011)
37. A. Taschin, P. Bartolini, R. Eramo, R. Righini, R. Torre, *Nat. Commun.* **4**, 2401 (2013)
38. K. Amann-Winkel et al., *Proc. Natl. Acad. Sci.* **110**, 17720 (2013)
39. K.H. Kim, A. Späh, H. Pathak et al., *Science* **358**, 1589–1593 (2017)

40. N. Stern, M. Seidl-Nigsch, T. Loerting, Proc. Natl. Acad. Sci. **116**, 9191–9196 (2019)
41. P.-L. Chau, A.J. Hardwick, Mol. Phys. **93**, 511 (1998)
42. J.R. Errington, P.G. Debenedetti, Nature **409**, 318 (2001)
43. I. Naberukhin Yu, V.P. Voloshin, N.N. Medvedev, Mol. Phys. **73**, 917 (1991)
44. A. Oleinikova, I. Brovchenko, J. Phys.: Condens. Matter **18**, S2247 (2006)
45. E. Shiratani, M. Sasai, J. Chem. Phys. **104**, 7671 (1996)
46. E. Shiratani, M. Sasai, J. Chem. Phys. **108**, 3264 (1998)
47. G.A. Appignanesi, J.A. Rodriguez Fris, F. Sciortino, Eur. Phys. J. E **29**, 305 (2009)
48. S.R. Accordino, J.A. Rodriguez Fris, F. Sciortino, G.A. Appignanesi, Eur. Phys. J. E **34**, 48 (2011)
49. D.C. Malaspina, J.A. Rodriguez Fris, G.A. Appignanesi, F. Sciortino, Eur. Phys. Lett. (EPL) **88**, 16003 (2009). <https://doi.org/10.1209/0295-5075/88/16003>
50. J.M.M. de Oca, J.A. Rodriguez Fris, S.R. Accordino, D.C. Malaspina, G.A. Appignanesi, Eur. Phys. J. E **39**, 124 (2016)
51. M.J. Cuthbertson, P.H. Poole, Phys. Rev. Lett. **106**, 115706 (2011)
52. R. Shi, H. Tanaka, J. Chem. Phys. **148**, 124503 (2018)
53. H. Tanaka, H. Tong, R. Shi, J. Russo, Revealing key structural features hidden in liquids and glasses, nature reviews. Physics **1**, 333–348 (2019). <https://doi.org/10.1038/s42254-019-0053-3>
54. J.M. Montes de Oca, F. Sciortino, G.A. Appignanesi, J. Chem. Phys. **152**, 244503 (2020). <https://doi.org/10.1063/5.0010895>
55. F. Sciortino, A. Geiger, H.E. Stanley, J. Chem. Phys. **96**, 3857 (1992)
56. F.H. Stillinger, *Energy Landscapes, Inherent Structures, and Condensed-Matter Phenomena* (Princeton University Press, Princeton, 2015)
57. F. Sciortino, Potential energy landscape description of supercooled liquids and glasses. J. Stat. Mech.: Theory Exp. **2005**, P05015 (2005)
58. P.G. Debenedetti, F.H. Stillinger, Nature **410**, 259 (2001)
59. A. Weber, F.H. Stillinger, J. Chem. Phys. **87**, 3252–3253 (1987)
60. H.J.C. Berendsen, J.R. Grigera, T.P. Stroatsma, J. Phys. Chem. **91**, 6269 (1987)
61. A.R. Verde, J.M. Montes, S.R. de Oca, L.M. Alarcón, Accordino, G.A. Appignanesi, J. Chem. Phys. **150**(24), 244504 (2019)
62. J.M.M. de Oca, S.R. Accordino, A.R. Verde, L.M. Alarcón, G.A. Appignanesi, Phys. Rev. E **99**(6), 062601 (2019)
63. R.G. Fernandez, J.L. Abascal, C. Vega, J. Chem. Phys. **124**, 144506 (2006)
64. K.T. Wikfeldt, A. Nilsson, L.G.M. Pettersson, Phys. Chem. Chem. Phys. **13**, 19918 (2011)
65. J. Gelman Constantin, A. Rodriguez Fris, G. Appignanesi, M. Carignano, I. Szleifer, H. Corti, Eur. Phys. J. E **34**, 126 (2011)
66. J.L.F. Abascal, E. Sanz, R.G. Fernandez, C. Vega, J. Chem. Phys. **122**, 4511 (2005)
67. J.L.F. Abascal, C. Vega, J. Chem. Phys. **123**, 234505 (2005)
68. S. Rick, J. Chem. Phys. **120**, 6085 (2004)
69. B. Santraa, R.A. DiDtasio Jr., F. Martelli, R. Car, Mol. Phys. **113**, 2829 (2015)
70. A.R. Verde, J.M.M. de Oca, S.R. Accordino, L.M. Alarcón, G.A. Appignanesi, Eur. Phys. J. E **44**(4), 1–10 (2021)
71. Shinji Saito, Biman Bagchi, Iwao Ohmine, J. Chem. Phys. **149**, 124504 (2018)
72. D. Laage, J.T. Hynes, Science **311**, 832 (2006)
73. J. Finney, A. Hallbrucker, I. Kohl, A. Soper, D. Bowron, Phys. Rev. Lett. **88**, 225503 (2002)
74. J. Škvára, F. Moučka, I. Nezbeda, J. Mol. Liquids **261**, 303 (2018)
75. V.F. Hernandez, M.S. Marques, J.R. Bordin, J. Phys.: Condens. Matter **34**, 024002 (2022)
76. F. Martelli, F. Leoni, F. Sciortino, J. Russo, J. Chem. Phys. **153**, 104503 (2020)
77. J. Carrasquilla, R. Melko, Nat. Phys. **13**, 431 (2017)
78. K. Swanson, S. Trivedi, J. Lequieu, K. Swanson, R. Kondorafg, Soft Matter **16**, 435 (2020)
79. E. Boattini, S. Marín-Aguilar, S. Mitra, G. Foffi, F. Smalenburg, L. Filion, Nat. Commun. **11**, 5479 (2020)
80. E. Boattini, M. Dijkstra, L. Filion, J. Chem. Phys. **151**, 154901 (2019)
81. P. Geiger, C. Dellago, J. Chem. Phys. **139**, 164105 (2013)
82. F. Martelli et al., Phys. Rev. B **97**, 064105 (2018)
83. F. Martelli, J. Crain, G. Franzese, ACS Nano **14**, 8616 (2020)

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.