Check for updates

OPEN

Topological nature of the liquid–liquid phase transition in tetrahedral liquids

Andreas Neophytou [ID][1 ✉], Dwaipayan Chakrabarti [ID][1 ✉] and Francesco Sciortino [ID][2 ✉]

The first-order phase transition between two tetrahedral networks of different density—introduced as a hypothesis to account for the anomalous behaviour of certain thermodynamic properties of deeply supercooled water—has received strong support from a growing body of work in recent years. Here we show that this liquid–liquid phase transition in tetrahedral networks can be described as a transition between an unentangled, low-density liquid and an entangled, high-density liquid, the latter containing an ensemble of topologically complex motifs. We first reveal this distinction in a rationally designed colloidal analogue of water. We show that this colloidal water model displays the well-known water thermodynamic anomalies as well as a liquid–liquid critical point. We then investigate water, employing two widely used molecular models, to demonstrate that there is also a clear topological distinction between its two supercooled liquid networks, thereby establishing the generality of this observation, which might have far-reaching implications for understanding liquid–liquid phase transitions in tetrahedral liquids.

A salient feature of liquid water is the anomalous behaviour of its thermodynamic response functions upon cooling, the most famous being the density maximum at ambient pressure[1–6]. In an effort to explain the origin of water's thermodynamic anomalies, enhanced in supercooled states, Poole et al. suggested the presence of a first-order liquid–liquid phase transition (LLPT) line in the supercooled region of the pressure–temperature (*P*–*T*) phase diagram[7]. In this scenario, this transition line separates two liquid phases formed of transient hydrogen-bond networks—the low-density liquid (LDL) and high-density liquid (HDL)—and terminates at a liquid–liquid critical point (LLCP)[7]. The anomalous behaviour of liquid water, along with water polyamorphism, was then suggested to be a manifestation of the presence of this LLPT. Recent numerical studies[8,9], as well as experiments[10–12] (despite the difficulties induced by the rapid formation of ice around the predicted critical temperature and pressure conditions), have provided strong support to this fascinating hypothesis.

Several order parameters, based on geometric or energetic criteria, have been proposed to characterize the transition[13–20]. In particular, the local structure index[13] and the r_5 parameter[14]—two order parameters that effectively measure the distance between the first and second coordination shells—have been routinely applied. However, along the years, it has become increasingly apparent that a proper description of the LLPT requires the development of order parameters which include information beyond that of the single-molecule neighbourhood. This led to the introduction of other order parameters, such as ζ (ref. [21]), V_4 (ref. [15]) and the total node communicability[20], which include some information on the connectivity properties of the hydrogen-bond network[19]. It has also become clear that network interpenetration[22–24] plays a key role in the LLPT in tetrahedral liquids[23,25–27], with the LDL being composed of a single random tetrahedral network and the HDL consisting of two (or more[23]) disordered but locally interpenetrating networks. In this context, a structure-based order parameter that fundamentally distinguishes the two liquids at the microscopic level will be important to develop our understanding of the mechanism for achieving a difference in density and, hence, the physical origin of the LLPT.

A colloidal model of water. Colloidal patchy particles, due to their capacity to form network liquids via encoded self-assembly information[25], combined with their synthetic availability[28–32], are ideally suited to investigating universal aspects of the LLPT, as well as potentially facilitating detailed experimental investigations at single-particle resolution. Therefore, guided by recent advances in programming their hierarchical self-assembly[33] and in understanding the conditions necessary for facilitating a LLPT—bond flexibility[25,26] and network interpenetration[23,26]—we rationally designed a colloidal analogue of water. Figure 1a schematically illustrates this 'colloidal water' model. We considered designer triblock patchy particles[28,30,34] with two distinct attractive patches that encode tetrahedrality—key to the uniqueness of water[5]. The tetrahedrality emerges as the triblock patchy particles undergo two-stage assembly upon cooling. The energetics and geometry of the patches, labelled A and B, encode the desired staged-assembly information into the particles[33,35,36]. The energetics are chosen to facilitate the formation of discrete tetrahedral clusters, initially driven by strong A–A interactions, and then a tetrahedral network fluid via weaker B–B interactions (Fig. 1a and Supplementary Fig. S1). The width of the A patches is judiciously chosen so as to favour the formation of a tetrahedral cluster fluid upon completion of the first stage of assembly[33,35,36]. The width of each B patch is instead chosen to favour its interaction with only one other B patch. The discrete tetrahedral clusters then behave as secondary building blocks, similar to tetrahedral patchy particles, that can interact with one another via B patches at lower temperatures to drive a second stage of assembly, leading to the formation of a liquid network. In this case, the patch widths are chosen not only to enforce the desired valency constraints but also to maximize the flexibility of the bonds to hinder the onset of crystallization[25]. As a result, the tetrahedral clusters formed by the triblock patchy particles can deviate from the ideal, tetrahedral symmetry. Note that the volume occupied by a tetrahedral cluster is less than that of a spherical patchy particle of equivalent size (Fig. 1a), thereby easing the formation of locally interpenetrating networks.

[1]School of Chemistry, University of Birmingham, Birmingham, UK. [2]Dipartimento di Fisica, Sapienza Università di Roma, Rome, Italy.
✉e-mail: axn241@student.bham.ac.uk; d.chakrabarti@bham.ac.uk; francesco.sciortino@uniroma1.it
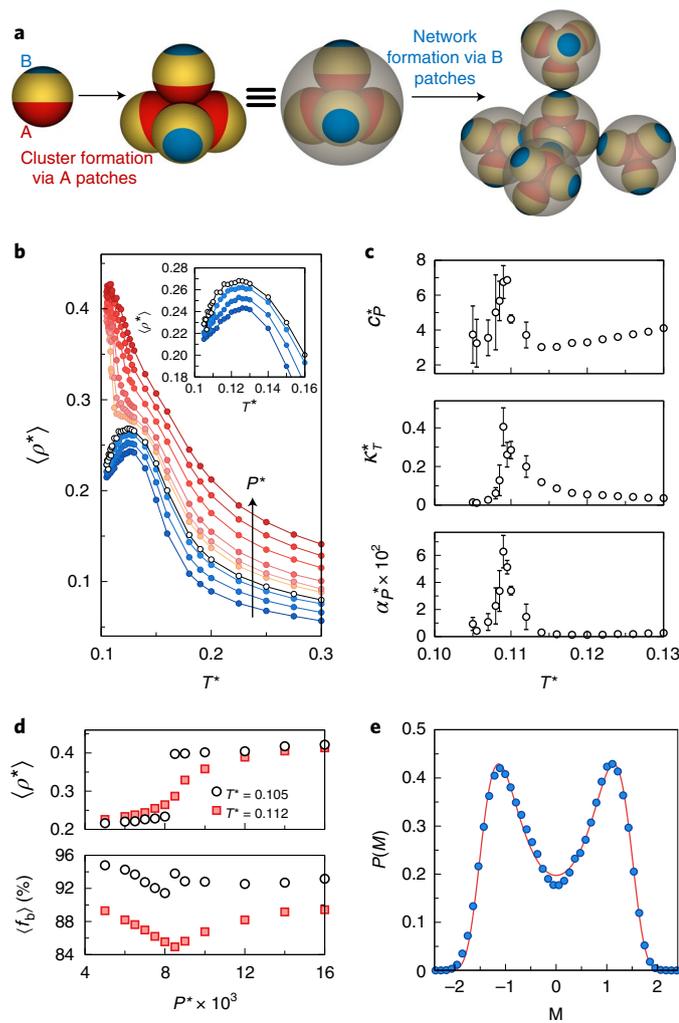
Fig. 1 | Thermodynamic anomalies and LLCP in colloidal water. a, Schematic of the hierarchical self-assembly of triblock patchy particles leading to a colloidal model of water. The two patches, labelled A and B, are of different sizes and form bonds of different strengths. The A patches (red) form stronger bonds than the B patches (blue) so as to encode two-stage assembly upon cooling. **b**, The evolution of the reduced density $\rho^\star$ as a function of the reduced temperature $T^\star$ for different reduced pressures $P^\star$, where $P^\star \times 10^3 = 5, 6, 7, 7.5, 8.5, 9, 10, 12, 14$ and 16. The arrow indicates the direction of increasing $P^\star$. The inset highlights the density maximum for $P^\star \times 10^3 = 5, 6, 7$ and 7.5. **c**, The evolution of the reduced thermal expansion coefficient ($\alpha_P^\star$), isothermal compressibility ($\kappa_T^\star$) and isobaric heat capacity ($c_P^\star$) as functions of $T^\star$ at $P^\star = 0.0085$ (close to the critical pressure). Error bars represent the standard deviation over five sets of Monte Carlo trajectories, each of 1×10^8 cycles. **d**, The dependence of $\rho^\star$ and the fraction of BB bonds formed (f_b) on $P^\star$ at $T^\star = 0.105$ and $T^\star = 0.112$. **e**, The distribution of the order parameter M for colloidal water (blue symbols), calculated using histogram reweighing, with $T^\star \approx 0.1075$, $P^\star \approx 0.0082$ and $s \approx 0.627$, compared with the corresponding 3D Ising universal distribution (solid red line).

Thermodynamic anomalies of colloidal water. To establish that the chosen system of triblock patchy particles can indeed be considered as a colloidal analogue of water, we first demonstrate that it captures water's thermodynamic anomalies. Figure 1b shows the reduced temperature ($T^\star \equiv k_B T / \epsilon_{BB}$, where k_B is the Boltzmann constant and ϵ_{BB} is the well depth of patch B–patch B interactions) dependence of the reduced density ($\rho^\star \equiv \rho \sigma^3$, where σ is the hard-sphere diameter of a triblock patchy particle) along different isobars as determined

by Monte Carlo simulations (Methods). Below a certain reduced pressure ($P^\star \equiv P\sigma^3/\epsilon_{BB}$), the colloidal water model displays a density maximum upon cooling. Upon further cooling, $\rho^\star$ approaches either a low or a high density value, depending on the pressure. Figure 1c shows the temperature dependence of the isobaric thermal expansion coefficient ($\alpha_P^\star$), isothermal compressibility ($\kappa_T^\star$) and isobaric heat capacity ($c_P^\star$) at a select pressure, highlighting that, upon cooling, colloidal water shows non-monotonic behaviour in its thermodynamic response functions.

Figure 1d shows the dependence of $\rho^\star$ and the fraction of bonds f_b (defined as the fraction of bonded B patches, a quantity directly reflective of the potential energy of the system) on $P^\star$. We find that, below a certain $T^\star$, there is a discontinuous jump in $\rho^\star$ and f_b as $P^\star$ changes, signalling a first-order phase transition. In contrast, $\rho^\star$ and f_b change continuously with $P^\star$ above this critical temperature, indicating that there is a LLCP. To locate the critical point, we followed a procedure similar to that applied recently in the study of a molecular model of water[9] to calculate the order parameter distribution $P(M)$, where $M = \rho^\star - sv^\star$ is a combination of the density, the potential energy per particle ($v^\star = \mathcal{V}/(N\epsilon_{BB})$, where $\mathcal{V}$ is the total potential energy) and a field-mixing parameter s (ref. [37]). Upon performing a non-linear fit to identify the critical temperature (T_c) and critical pressure (P_c), we find that $P(M)$ closely matches the distribution of the magnetization at the critical point of the three-dimensional (3D) Ising model (Fig. 1e), confirming the presence of an LLCP in colloidal water.

Topological nature of the LLPT in colloidal water. Topological concepts have become central in the description of physical[38–40], biological[41,42], social[43] and financial[44,45] networks. The topological properties of the interaction matrix,[39] owing to their unexpected connections with hidden conservation laws, have been shown to affect the response of the system to external perturbation and to play a key role in phase transitions. In this context, a pertinent question is whether the LLPT also has a topological character. Therefore, to identify any topological features, we inspected configurations of the LDL and HDL, finding that linked and knotted ring structures were noticeably present in the networks of the HDL phase but seemingly absent from the LDL phase. Figure 2a,b shows representative snapshots of the two liquid phases, where examples of these topologically complex motifs in the HDL networks are highlighted: a trefoil knot and a Hopf link[46] (Supplementary Video 1). In addition to knots and links, we found a large number of knotted theta curves[47,48] in the HDL networks, which can be viewed as two entangled fused rings (that is, rings that share at least two vertices). Examples of these knotted theta curves are shown in Fig. 2c and Supplementary Videos 2–4.

To investigate the importance of such topological features in the LLPT, we quantified the degree of 'entanglement' in the LDL and HDL networks using the concept of helicity. The helicity has previously been used to provide a measure of the topological properties of vector fields, including magnetic fields[49] and vortex fields in fluids[50]. This scalar quantity measures the amount of entwining in the set of closed curves contained within a system. For the particle-based networks considered here, these closed curves correspond to the set of rings formed by bonded particles. Each pair of disjoint (and therefore distinct) rings R_i and R_j contributes to the helicity with a term proportional, in essence, to how many times the two rings wind around one another (alternatively, this can be thought of as the number of times ring R_i pierces the surface bounded by ring R_j). Also, each individual ring contributes with a term proportional to how many times it loops around itself. Formally, the helicity is defined as a double sum over the total number of rings $\mathcal{N}_R$ of the double line integral, known as the Gauss linking integral:

$$\mathcal{H} = \sum_{i=1}^{\mathcal{N}_R} \sum_{j \geq i}^{\mathcal{N}_R} \frac{1}{4\pi} \oint_{R_i} \oint_{R_j} \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3} \cdot (\mathrm{d}\mathbf{r}_j \times \mathrm{d}\mathbf{r}_i), \quad (1)$$
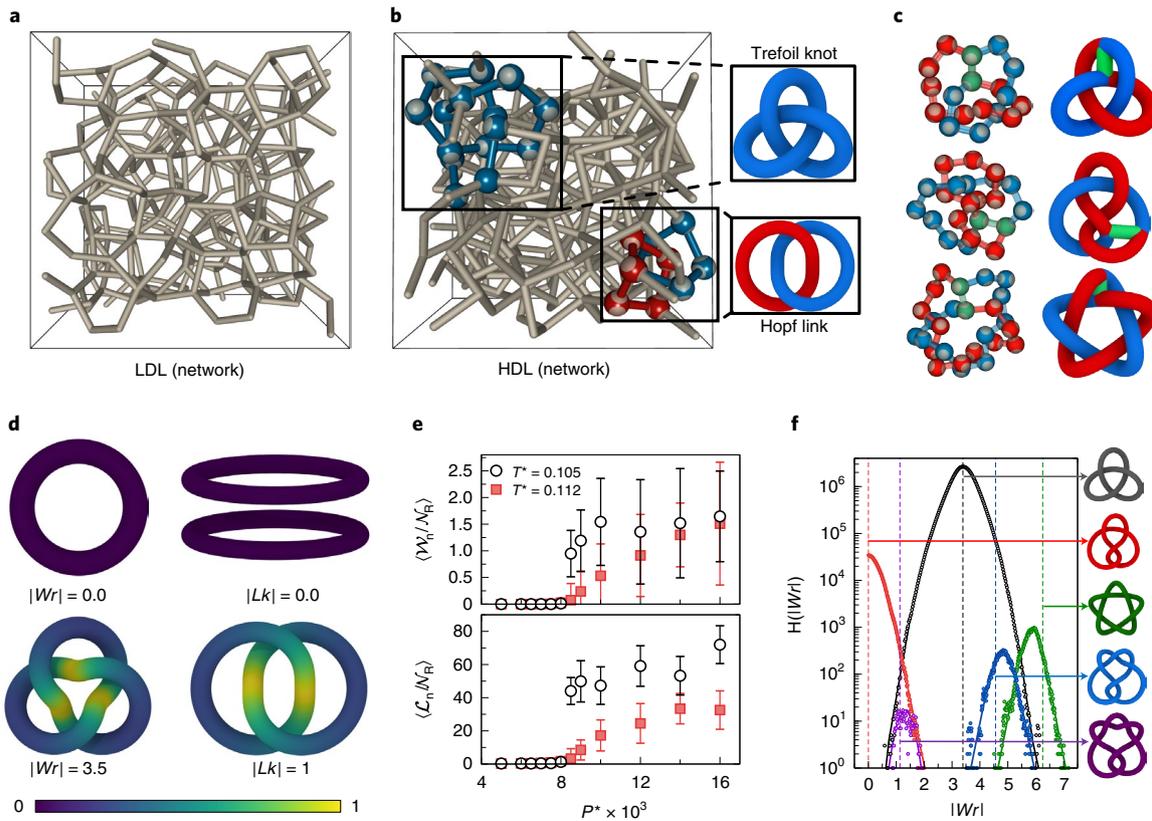
Fig. 2 | Characterization of the LLPT in colloidal water by identifying links and knots. a,b, Representative snapshots of the LDL (**a**) and HDL (**b**) networks of colloidal water at $T^* = 0.105$, a temperature below the critical temperature T_c, and pressures either side of the critical pressure P_c ($P^* = 0.005$ and $P^* = 0.016$). Here, vertices correspond to the centres of tetrahedral clusters, while edges indicate existing B–B bonds. Tetrahedral clusters in the HDL network forming a trefoil knot (top) and a Hopf link (bottom) are highlighted in **b**, visualized in a reduced representation as tetrahedral patchy particles. **c,** Illustrative snapshots of theta curves found in the HDL networks and their corresponding idealized representations. **d,** The linking number (Lk) or writhe (Wr) for an idealized unknot (top left), unlink (top right), trefoil knot (bottom left) and Hopf link (bottom right). Each point in the motifs is coloured according to the corresponding (relative) contribution to the Gauss linking integral (equation (1)). **e,** The dependence of the average network linking number ($\mathcal{L}_n$) and the network writhe ($\mathcal{W}_n$) divided by the total number of rings in the network ($\mathcal{N}_R$) as a function of P^* at $T^* = 0.105$ ($T < T_c$) and $T^* = 0.112$ ($T > T_c$). $\mathcal{L}_n$ and $\mathcal{W}_n$ were computed using rings of size up to $l_{max} = 13$. Error bars represent the standard deviation over two Monte Carlo trajectories, each of 4×10^7 cycles. **f,** Frequency distribution for the writhe (Wr), separated for different unique knotted structures observed in the HDL network, at $T^* = 0.105$ and $P^* = 0.016$ with $l_{max} = 15$. The idealized knot associated with each distribution is colour coded and shown to the right of the plot. The dashed vertical lines indicate the writhe values of the idealized knots[57], and the solid lines through the data points are a guide to the eye.

where $\mathbf{r}_i$ and $\mathbf{r}_j$ are points on the two rings R_i and R_j, respectively. In equation (1), $d\mathbf{r}_i$ and $d\mathbf{r}_j$ are infinitesimal vectors tangential to the two rings R_i and R_j located at points $\mathbf{r}_i$ and $\mathbf{r}_j$, respectively. In the case where $i \neq j$, the Gauss linking integral gives the linking number (Lk_{ij}) of the two curves and takes on only integer values, whereas for $i = j$, it corresponds to the writhe (Wr_i) of a single curve and can take on any real value. Equation (1) can therefore be rewritten as the sum of the linking number and the writhe as

$$\mathcal{H} = \sum_{i>j}^{\mathcal{N}_R} Lk_{ij} + \sum_{i=1}^{\mathcal{N}_R} Wr_i. \tag{2}$$

A visual representation of the Gauss linking integral for a set of closed loops, along with their corresponding writhe or linking number, is shown in Fig. 2d.

To establish the prevalence of topologically complex motifs in a liquid network, we adapted the concept of the helicity to define two quantities: the network linking number[39] ($\mathcal{L}_n$) and the network writhe ($\mathcal{W}_n$). To emphasize the degree of entanglement in a network, $\mathcal{L}_n$ and $\mathcal{W}_n$ are computed by summing up the absolute values of Lk_{ij} and Wr_i, respectively, to ignore the intrinsic orientation of the

closed curves (in other words, we ignore the chirality of the topological objects). Additionally, the writhe is only computed for the set $\mathcal{N}_k$ of knotted objects that includes knotted rings and pairs of fused rings that contain a knotted path. Including the latter in $\mathcal{N}_k$ ensures that the contributions of theta curves (such as those shown in Fig. 2c) are captured by $\mathcal{W}_n$. We formally define $\mathcal{L}_n$ and $\mathcal{W}_n$ as

$$\mathcal{L}_n = \sum_{i>j}^{\mathcal{N}_R} |Lk_{ij}| \quad \text{and} \quad \mathcal{W}_n = \sum_{i=1}^{\mathcal{N}_k} |Wr_i| \tag{3}$$

We also note that, in some instances, the helicity can include a twist term when the structures under consideration are constructed from bundles of closed curves (as is the case for supercoiled DNA). However, since this is not the case with the bond rings considered here, the twist can practically be ignored.

We computed $\mathcal{L}_n$ and $\mathcal{W}_n$ by first identifying all the rings in the system up to a maximum size of l_{max} (which defines the maximum number of particles forming a ring) and then calculating either (1) the linking number for all disjoint rings or (2) the writhe for each ring that is knotted and for entangled fused rings (such as the knotted theta curves shown in Fig. 2c). This then becomes akin to
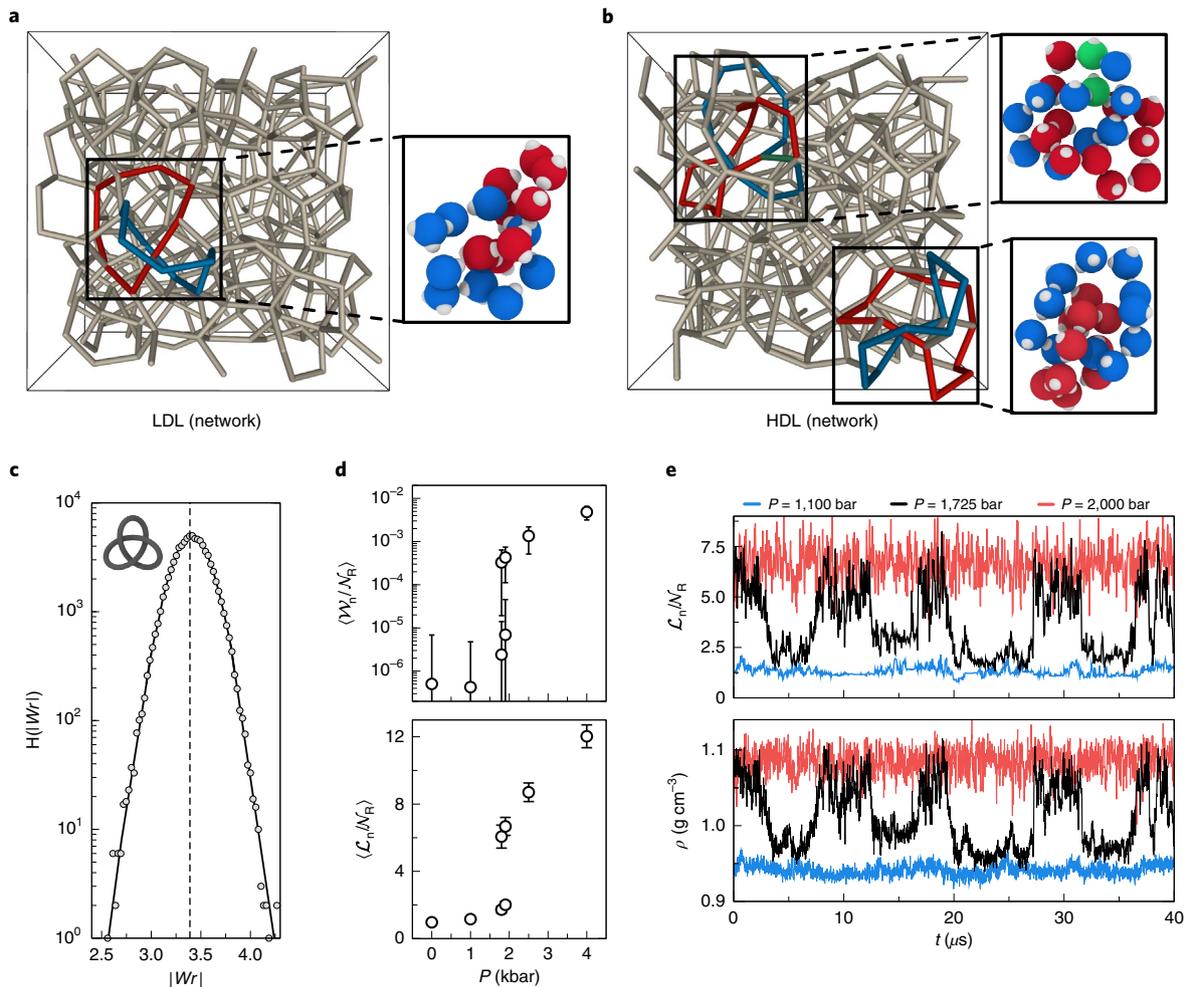
Fig. 3 | Characterization of the LLPT in molecular water by identifying links and knots. a,b, Illustrative snapshots of the LDL (**a**) and HDL (**b**) hydrogen-bond networks of molecular water, simulated using the TIP4P/Ice potential, at $T=188\,K$ and $P=1,100\,bar$ and $P=2,000\,bar$, respectively. This temperature is slightly below the critical temperature ($T_c=188.8\,K$ (ref. [9])), and the pressures are below and above the critical pressure ($P_c=1,725\,bar$ (ref. [9])). In the LDL, only a few links are present (one is highlighted in **a**). In the HDL, both linked and knotted motifs are present (highlighted in **b**). **c,** The frequency distribution of the writhe (Wr) of the knotted paths found in the HDL network (essentially all being trefoil knots) at $T=188\,K$ and $P=2,000\,bar$. The inset shows the idealized configuration of a trefoil knot, and the dotted vertical line indicates the value of the writhe for the idealized trefoil knot[57]. **d,** Dependence of $\mathcal{L}_n$ and $\mathcal{W}_n$, computed using rings of sizes up to $l_{max}=13$, divided by $\mathcal{N}_R$, as a function of P at $T=188\,K$ for $N=1,000$ molecules. Error bars represent the standard deviation as calculated along the corresponding molecular dynamics trajectory. **e,** The fluctuations in the density (ρ) and $\mathcal{L}_n$ (computed using $l_{max}=13$) with time (t) along an isobaric–isothermal molecular dynamics trajectory for $N=300$ molecules of TIP4P/Ice water at a temperature of $T=188\,K$ for $P=1,100\,bar$, $P=1,725\,bar$ and $P=2,000\,bar$.

identifying linked and knotted paths in the spatial embedding of a graph[48,51] (full details on how $\mathcal{L}_n$ and $\mathcal{W}_n$ are computed in practice are given in Methods). Because we only identified links between rings of sizes $l \leq l_{max}$ and knotted structures of sizes $l \leq (2l_{max} - 2)$, the computed values of $\mathcal{L}_n$ and $\mathcal{W}_n$ depend on the chosen l_{max}. The LDL and HDL networks were found to show asymptotic limiting behaviour in their topological properties as a function of l_{max}, above a critical value (Supplementary Figs. S2 and S3). Therefore, we computed $\mathcal{L}_n$ and $\mathcal{W}_n$ using the smallest l_{max} value that was found to be in the asymptotic regions of both the LDL and HDL phases ($l_{max} \geq 13$). Figure 2e shows the average values of $\mathcal{L}_n$ and $\mathcal{W}_n$ as a function of pressure for $T < T_c$ and $T > T_c$ with $l_{max} = 13$. Both $\mathcal{L}_n$ and $\mathcal{W}_n$ display a discontinuous change with pressure for sub-critical temperatures ($T < T_c$), concomitant with the discontinuity in the density (Fig. 1d), while they change continuously for super-critical temperatures.

The transition between the LDL and HDL phases is clearly captured by $\mathcal{L}_n$ when $l_{max} \geq 5$ (Extended Data Fig. 1). However, as knots require longer chains of particles to form, they require a larger l_{max} to be identified, and so, $\mathcal{W}_n$ only shows a clear discontinuity when $l_{max} \geq 10$ (Extended Data Fig. 1). Additionally, this means that there are fewer knots than links in the network, which in turn, means that $\mathcal{W}_n$ is smaller than $\mathcal{L}_n$ for a given l_{max}. Yet, there are a variety of knotted objects in the HDL networks that contribute to the nonzero value of $\mathcal{W}_n$. Figure 2f shows the frequency distribution of the writhe for different knotted objects found in the HDL network with $l_{max} = 15$. The majority of these structures contain trefoil knots, but we also find figure-of-eight, cinquefoil, three-twist and Stevedore knots[46]. Similar to randomly generated knotted rings, the ensemble average of the writhe for each of these knots correlates well with the writhe of the corresponding idealized configuration (an idealized knot configuration is one which allows the volume-to-surface area ratio to be maximized[52]).

Topological nature of the LLPT in molecular water. It is clear that both $\mathcal{L}_n$ and $\mathcal{W}_n$ can serve as a topological order parameter to describe the LLPT in colloidal water, providing strong evidence that the LLPT is a transition from an 'unentangled' to an 'entangled' network. This raises the question of whether entanglement also underpins the LLPT in supercooled water. Recently, it was shown that both the TIP4P/Ice[53] and TIP4P/2005[54] molecular models of water, which do not explicitly enforce tetrahedral coordination, exhibit a first-order LLPT and corresponding LLCP[9]. We therefore performed a similar topological analysis for TIP4P/Ice and TIP4P/2005. Figure 3a,b shows representative snapshots of the hydrogen-bond networks for the LDL and HDL of TIP4P/Ice. We find that the LDL contains a small number of links and practically no knotted structures (with only one or two knots occasionally appearing in a frame), while the HDL contains trefoil knots and theta curves in addition to links (which are also present in more abundance than in the LDL). Figure 3c shows that the ensemble average of the writhe for these trefoil motifs in the HDL of TIP4P/Ice also correlates well with the writhe of the idealized trefoil knot, as was the case for colloidal water. Additionally, as for colloidal water, Fig. 3d shows that, for TIP4P/Ice, both $\mathcal{L}_n$ and $\mathcal{W}_n$ jump at the critical pressure, suggesting that entanglement also underpins the LLPT in molecular water.

A universal property of systems close to a second-order critical point (when $T \approx T_c$ and $P \approx P_c$) is wide oscillations in relevant order parameters, as configurations belonging to both phases are sampled along the trajectory[37]. To check whether the topological order parameter $\mathcal{L}_n$ does indeed properly distinguish between the two liquid phases of the TIP4P/Ice and TIP4P/2005 models, we calculated $\mathcal{L}_n$ along 40 μs molecular dynamics trajectories of supercooled water for different values of P close to T_c. Figure 3e shows the evolution of ρ and $\mathcal{L}_n$ along the molecular dynamics trajectory for a system of $N=300$ TIP4P/Ice molecules at three different pressures. Above and below the critical pressure, $\mathcal{L}_n$ displays quite different values, indicating a notable topological difference between the two liquids. At the critical pressure, wide fluctuations in $\mathcal{L}_n$ are observed, closely correlating with those of ρ. Extended Data Fig. 2 shows that the TIP4P/2005 model also displays these wide fluctuations in $\mathcal{L}_n$ and ρ when $P \approx P_c$, confirming that the LLPT can indeed be interpreted as a topological transition.

Conclusions. We have designed a colloidal analogue of water—a tetrahedral network liquid self-assembled from designer triblock patchy colloidal particles via tetrahedral clusters. This colloidal water model captures the anomalous thermodynamic behaviour of supercooled water, including the well-known density maximum, and displays two structurally different network liquid phases. With the synthetic feasibility of the triblock patchy particles under consideration[30], colloidal water presents itself as an ideal model system, amenable to experimental investigation at the single-particle level, for gleaning fundamental understanding of LLPTs in tetrahedral liquids.

By analysing configurations from this colloidal model and from two widely used molecular models[53] for water, we have established that the LDL and HDL networks associated with the LLPT in all three cases are topologically distinguishable. The LDL in all these cases is practically an 'unentangled network', consisting primarily of unknots, while the HDL is an 'entangled network' containing an ensemble of topologically non-trivial motifs. For colloidal as well as molecular models, $\mathcal{L}_n$ and $\mathcal{W}_n$—metrics able to capture the degree of entanglement—act as suitable order parameters, highlighting that the LLPT can also be interpreted as a topological phase transition.

Uncovering the topological distinction between the LDL and HDL networks allows us to understand the physical mechanism underpinning the LLPT from a new perspective that sheds light on its microscopic origin. The LLPT occurs at low temperatures at which the system tends to form a liquid network with the maximum

number of bonds possible. At low pressures, it is well established that this is achieved by forming an open random tetrahedral network. However, as the pressure increases, the system tends also to minimize its volume. In tetrahedral liquids, associated with this drive to maximize the number of bonds, there is an increase in the number of rings in the network. Due to the directionality of the interactions, the system cannot minimize its volume by simply deforming the rings while preserving the network connectivity because there is a limit to the flexibility of the bonds. Instead, the system is able to simultaneously minimize its volume and maximize the number of bonds in the network by forming knots and links. Additionally, we note that, unlike knots and links formed by covalently bonded chains or networks[40], the LDL and HDL are transient networks (that is, bonds are constantly breaking and forming) and so their topological state is not constant. Despite this evolution, the topological signatures for each phase are properly defined.

The helicity has been used previously to study knotted vortices in classical fluids[55] and superfluids[56], where these knots were found to untie via a universal topological mechanism. It would be of interest to use the metrics for the degree of entanglement used here to follow the nucleation of the LDL out of the HDL, and vice versa, to see whether this universal topological mechanism is also relevant in the context of the LLPT. Finally, owing to their topological complexity, these entangled liquids should possess physical properties[38,40] distinct from simple liquids, which will merit exploration.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-022-01698-6.

References

1. Speedy, R. & Angell, C. Isothermal compressibility of supercooled water and evidence for a thermodynamic singularity at −45°C. *J. Chem. Phys.* **65**, 851–858 (1976).
2. Angell, C., Sichina, W. & Oguni, M. Heat capacity of water at extremes of supercooling and superheating. *J. Phys. Chem.* **86**, 998–1002 (1982).
3. Angell, C. A. Supercooled water. *Annu. Rev. Phys. Chem.* **34**, 593–630 (1983).
4. Debenedetti, P. G. Supercooled and glassy water. *J. Phys. Condens. Matter* **15**, R1669 (2003).
5. Russo, J., Akahane, K. & Tanaka, H. Water-like anomalies as a function of tetrahedrality. *Proc. Natl Acad. Sci. U. S. A.* **115**, E3333–E3341 (2018).
6. Russo, J., Leoni, F., Martelli, F. & Sciortino, F. The physics of empty liquids: from patchy particles to water. *Rep. Prog. Phys.* **85**, 016601 (2021).
7. Poole, P. H., Sciortino, F., Essmann, U. & Stanley, H. E. Phase behaviour of metastable water. *Nature* **360**, 324–328 (1992).
8. Palmer, J. C. et al. Metastable liquid–liquid transition in a molecular model of water. *Nature* **510**, 385–388 (2014).
9. Debenedetti, P. G., Sciortino, F. & Zerze, G. H. Second critical point in two realistic models of water. *Science* **369**, 289–292 (2020).
10. Kim, K. H. et al. Experimental observation of the liquid-liquid transition in bulk supercooled water under pressure. *Science* **370**, 978–982 (2020).
11. Kim, K. H. et al. Maxima in the thermodynamic response and correlation functions of deeply supercooled water. *Science* **358**, 1589–1593 (2017).
12. Perakis, F. et al. Diffusive dynamics during the high-to-low density transition in amorphous ice. *Proc. Natl Acad. Sci. U. S. A.* **114**, 8193–8198 (2017).
13. Shiratani, E. & Sasai, M. Molecular scale precursor of the liquid–liquid phase transition of water. *J. Chem. Phys.* **108**, 3264–3276 (1998).
14. Cuthbertson, M. J. & Poole, P. H. Mixturelike behavior near a liquid-liquid phase transition in simulations of supercooled water. *Phys. Rev. Lett.* **106**, 115706 (2011).
15. de Oca, J. M. M., Sciortino, F. & Appignanesi, G. A. A structural indicator for water built upon potential energy considerations. *J. Chem. Phys.* **152**, 244503 (2020).
16. Tanaka, H., Tong, H., Shi, R. & Russo, J. Revealing key structural features hidden in liquids and glasses. *Nat. Rev. Phys.* **1**, 333–348 (2019).

17. Tanaka, H. Liquid–liquid transition and polyamorphism. *J. Chem. Phys.* **153**, 130901 (2020).

18. Steinberg, L., Russo, J. & Frey, J. A new topological descriptor for water network structure. *J. Cheminform.* **11**, 1–11 (2019).

19. Foffi, R., Russo, J. & Sciortino, F. Structural and topological changes across the liquid–liquid transition in water. *J. Chem. Phys.* **154**, 184506 (2021).

20. Faccio, C., Benzi, M., Zanetti-Polzi, L. & Daidone, I. Low- and high-density forms of liquid water revealed by a new medium-range order descriptor. *J. Mol. Liq.* **355**, 118922 (2022).

21. Russo, J. & Tanaka, H. Understanding water's anomalies with locally favoured structures. *Nat. Commun.* **5**, 1–11 (2014).

22. Kell, G. S. Completely hydrogen-bonded interpenetrating networks in low-temperature liquid water: their geometrical possibility and stability. *Chem. Phys. Lett.* **30**, 223–226 (1975).

23. Hsu, C. W., Largo, J., Sciortino, F. & Starr, F. W. Hierarchies of networked phases induced by multiple liquid-liquid critical points. *Proc. Natl Acad. Sci. U. S. A.* **105**, 13711–13715 (2008).

24. Hsu, C. W. & Starr, F. W. Interpenetration as a mechanism for liquid–liquid phase transitions. *Phys. Rev. E* **79**, 041502 (2009).

25. Smallenburg, F. & Sciortino, F. Liquids more stable than crystals in particles with limited valence and flexible bonds. *Nat. Phys.* **9**, 554–558 (2013).

26. Smallenburg, F., Filion, L. & Sciortino, F. Erasing no-man's land by thermodynamically stabilizing the liquid–liquid transition in tetrahedral particles. *Nat. Phys.* **10**, 653–657 (2014).

27. Ciarella, S., Gang, O. & Sciortino, F. Toward the observation of a liquid-liquid phase transition in patchy origami tetrahedra: a numerical study. *Eur. Phys. J. E* **39**, 1–6 (2016).

28. Pawar, A. B. & Kretzschmar, I. Patchy particles by glancing angle deposition. *Langmuir* **24**, 355–358 (2008).

29. Wang, Y. et al. Colloids with valence and specific directional bonding. *Nature* **491**, 51–55 (2012).

30. Chen, Q., Bae, S. C. & Granick, S. Staged self-assembly of colloidal metastructures. *J. Am. Chem. Soc.* **134**, 11080–11083 (2012).

31. Gong, Z., Hueckel, T., Yi, G.-R. & Sacanna, S. Patchy particles made by colloidal fusion. *Nature* **550**, 234–238 (2017).

32. Hueckel, T., Hocky, G. M. & Sacanna, S. Total synthesis of colloidal matter. *Nat. Rev. Mater.* **6**, 1053–1069 (2021).

33. Morphew, D., Shaw, J., Avins, C. & Chakrabarti, D. Programming hierarchical self-assembly of patchy particles into colloidal crystals via colloidal molecules. *ACS Nano* **12**, 2355–2364 (2018).

34. Chen, Q., Bae, S. C. & Granick, S. Directed self-assembly of a colloidal kagome lattice. *Nature* **469**, 381–384 (2011).

35. Rao, A. B. et al. Leveraging hierarchical self-assembly pathways for realizing colloidal photonic crystals. *ACS Nano* **14**, 5348–5349 (2020).

36. Neophytou, A., Manoharan, V. N. & Chakrabarti, D. Self-assembly of patchy colloidal rods into photonic crystals robust to stacking faults. *ACS Nano* **15**, 2668–2678 (2021).

37. Wilding, N. & Bruce, A. Density fluctuations and field mixing in the critical fluid. *J. Phys. Condens. Matter* **4**, 3087–3108 (1992).

38. Horner, K. E., Miller, M. A., Steed, J. W. & Sutcliffe, P. M. Knot theory in modern chemistry. *Chem. Soc. Rev.* **45**, 6432–6448 (2016).

39. Liu, Y., Dehmamy, N. & Barabási, A.-L. Isotopy and energy of physical networks. *Nat. Phys.* **17**, 216–222 (2021).

40. Orlandini, E. & Micheletti, C. Topological and physical links in soft matter systems. *J. Phys. Condens. Matter* **34**, 013002 (2021).

41. Patil, K. R. & Nielsen, J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl Acad. Sci. U. S. A.* **102**, 2685–2689 (2005).

42. Gross, T., D'Lima, C. J. D. & Blasius, B. Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.* **96**, 208701 (2006).

43. Centola, D. The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197 (2010).

44. Bonanno, G., Caldarelli, G., Lillo, F. & Mantegna, R. N. Topology of correlation-based minimal spanning trees in real and model markets. *Phys. Rev. E* **68**, 046130 (2003).

45. Diebold, F. X. & Yilmaz, K. On the network topology of variance decompositions: measuring the connectedness of financial firms. *J. Econom.* **182**, 119–134 (2014).

46. Rolfsen, D. *Knots and Links*, vol. 346 (American Mathematical Society, 2003).

47. Wolcott, K. in *Geometry and Topology* (eds McCrory, C. & Shifrin, T.) (Marcel Dekker, 1987).

48. Kauffman, L. H. Invariants of graphs in three-space. *Trans. Am. Math. Soc.* **311**, 697–710 (1989).

49. Berger, M. A. Introduction to magnetic helicity. *Plasma Phys. Control. Fusion* **41**, B167 (1999).

50. Moffatt, H. K. The degree of knottedness of tangled vortex lines. *J. Fluid Mech.* **35**, 117–129 (1969).

51. Conway, J. H. & McA. Gordon, C. Knots and links in spatial graphs. *J. Graph Theory* **7**, 445–453 (1983).

52. Katritch, V. et al. Geometry and physics of knots. *Nature* **384**, 142–145 (1996).

53. Abascal, J., Sanz, E., García Fernández, R. & Vega, C. A potential model for the study of ices and amorphous water: TIP4P/Ice. *J. Chem. Phys.* **122**, 234511 (2005).

54. Abascal, J. L. & Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **123**, 234505 (2005).

55. Kleckner, D. & Irvine, W. T. Creation and dynamics of knotted vortices. *Nat. Phys.* **9**, 253–258 (2013).

56. Kleckner, D., Kauffman, L. H. & Irvine, W. T. How superfluid vortex knots untie. *Nat. Phys.* **12**, 650–655 (2016).

57. Cerf, C. & Stasiak, A. A topological invariant to predict the three-dimensional writhe of ideal configurations of knots and links. *Proc. Natl Acad. Sci. U.S.A.* **97**, 3795–3798 (2000).

Methods

Colloidal water model. We employed the Kern–Frenkel pair potential[58,59], which has been used extensively to study patchy particles. Each particle, with a hard-core diameter σ, has (at opposite poles) two distinct, circular attractive patches, labelled A and B, having half-angles of θ_A and θ_B, respectively. The effective potential for a pair of patchy particles v_{ij} is

$$v_{ij}(\mathbf{r}_{ij}, \mathbf{\Omega}_i, \mathbf{\Omega}_j) = v_{ij}^{hs}(r_{ij}) + \sum_{\alpha,\beta=\{A,B\}} v_{\alpha\beta}^{sw}(r_{ij}) f(\mathbf{r}_{ij}, \hat{\mathbf{n}}_i^\alpha, \hat{\mathbf{n}}_j^\beta), \quad (4)$$

where $r_{ij} = |\mathbf{r}_{ij}|$ is the centre-to-centre distance between particles i and j and $\mathbf{\Omega}_i$ and $\mathbf{\Omega}_j$ describe the orientations of particles i and j, respectively. In equation (4), v_{ij}^{hs} is the hard-sphere pair potential, given by

$$v_{ij}^{hs}(r_{ij}) = \begin{cases} \infty & \text{if } r_{ij} < \sigma \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

and v_{ij}^{sw} is a square-well potential, given by

$$v_{\alpha\beta}^{sw}(r_{ij}) = \begin{cases} -\varepsilon_{\alpha\beta} & \text{if } \sigma \leq r_{ij} \leq (1 + \delta_{\alpha\beta})\sigma \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

where $\varepsilon_{\alpha\beta}$ and $\delta_{\alpha\beta}$ control the depth and range, respectively, of patch α–patch β attraction. The factor $f(\mathbf{r}_{ij}, \hat{\mathbf{n}}_i^\alpha, \hat{\mathbf{n}}_j^\beta)$ controls the angular dependence of the interaction between two patches and is given by

$$f(\mathbf{r}_{ij}, \hat{\mathbf{n}}_i^\alpha, \hat{\mathbf{n}}_j^\beta) = \begin{cases} 1 & \text{if } \hat{\mathbf{n}}_i^\alpha \cdot \hat{\mathbf{r}}_{ij} > \cos\theta_\alpha \text{ and } \hat{\mathbf{n}}_j^\beta \cdot \hat{\mathbf{r}}_{ji} > \cos\theta_\beta \\ 0 & \text{otherwise} \end{cases}, \quad (7)$$

where $\hat{\mathbf{n}}_i^\alpha$ is a normalized vector from the centre of particle i in the direction of the centre of patch α on its surface, thus depending on $\mathbf{\Omega}_i$. Similarly, $\hat{\mathbf{n}}_j^\beta$ is a normalized vector from the centre of particle j in the direction of the centre of patch β, thus depending on $\mathbf{\Omega}_j$. The total potential energy of the system was calculated by summing over the contributions from all distinct pairs ($\mathcal{V} = \sum_{i \neq j} v_{ij}$).

We employed $\varepsilon_{AA} = 5\varepsilon_{BB}$ along with $\varepsilon_{AB} = 0$ to encode a two-stage self-assembly process upon cooling[33]. We selected $\theta_A = 50°$ and $\theta_B = 26°$ because this choice, combined with the interaction strengths, ensures that an A patch interacts with three other A patches (to yield discrete tetrahedral clusters at relatively high temperatures) while a B patch interacts with only one other B patch (to bring the tetrahedral clusters together to form a tetrahedral network liquid at lower temperatures). The choice of $\varepsilon_{AB} = 0$, which is realizable via DNA-mediated interactions, facilitates staged assembly and allows for the use of a wider A patch without compromising the formation of self-limiting tetrahedral clusters at intermediate temperatures. Note that the use of an A patch with $\theta_A = 50°$ makes the tetrahedral clusters flexible enough to hinder crystallization from the tetrahedral liquids formed at lower temperatures. The ranges of the interactions are set by $\delta_{AA} = 0.05$ and $\delta_{BB} = 0.2$ (the value of δ_{AB} being irrelevant since $\varepsilon_{AB} = 0$). The longer range of B–B interactions facilitates the formation of a tetrahedral network liquid at low temperatures. In the present study, we used reduced units: the length in units of σ, the energy in units of ε_{BB} and the temperature in units of ε_{BB}/k_B, with the Boltzmann constant set to $k_B = 1$.

Monte Carlo simulations of colloidal water. All the Monte Carlo simulations were carried out with systems of $N = 1{,}000$ triblock patchy particles contained in a cubic box under periodic boundary conditions, using the minimum image convention. Each triblock patchy particle was treated as a rigid body whose orientational degrees of freedom were represented by quaternions. The potential energy was calculated using a spherical cutoff of $\sigma(1 + \delta_{BB})$, and a cell list was used for efficiency[60]. To enhance sampling in systems composed of clusters, we employed a combination of translational and rotational single-particle and cluster moves, where the acceptance probability for a translational or rotational cluster move was taken as[60]

$$\text{acc}(o \rightarrow n) = \min\left\{1, \exp\left[-\beta\left(\mathcal{V}(n) - \mathcal{V}(o)\right)\right] \prod_{i,j} \frac{1 - P_n(i,j)}{1 - P_o(i,j)}\right\} \quad (8)$$

where $\beta = 1/k_B T$, $\mathcal{V}(o)$ and $\mathcal{V}(n)$ correspond to the potential energy of the system before and after the cluster move, respectively, and $P(i,j)$ represents the probability that particles i and j share a bond, that is, that the patches interact. We defined a cycle as the number of attempts, equal to the number of particles N, to move (rotate or translate) a particle or a cluster (with equal probability). Additionally, in isothermal-isobaric (NPT) simulations, we included one volume move, on average, per cycle. We performed cluster volume moves in the logarithmic space of the volume (that is, $\ln(V_n) = \ln(V_o) + \Delta V$), where each cluster's centre of mass was scaled isotropically. Volume moves were accepted with probability[61]

$$\text{acc}(V_o \rightarrow V_n) = \min\left\{1, \exp\left[(N_{clstr} + 1) \ln\left(\frac{V_n}{V_o}\right)\right.\right.$$
$$\left.\left. - \beta P \Delta V - \beta \Delta \mathcal{V}\right] \prod_{i,j} \frac{1 - P_n(i,j)}{1 - P_o(i,j)}\right\}, \quad (9)$$

where N_{clstr} is the number of discrete clusters identified in the system. In the context of an anticipated two-stage assembly of the patchy colloidal particles via a discrete set of clusters, formed by A–A bonds, we constructed a simple rule to move the discrete clusters. We took $P(i,j) = 1$ if two particles share an A–A bond, and $P(i,j) = 0$ otherwise. Therefore, to ensure detailed balance is satisfied, any proposed cluster or volume moves that created a new A–A bond between particles i and j were rejected.

At low temperatures ($T^\star \leq 0.12$), two sets of simulations were carried out over the same range of state points, starting from a tetrahedral cluster fluid of triblock patchy particles, where the initial density was set to $\rho^\star = 0.2$ for one set of simulations and to $\rho^\star = 0.4$ for the other set. For each of these simulations, 3×10^8 Monte Carlo cycles were performed to ensure that equilibrium was attained. At higher temperatures, 1×10^8 Monte Carlo cycles were performed, starting with a density of $\rho^\star = 0.4$.

From the NPT simulations, we calculated the reduced isobaric thermal expansion coefficient ($\alpha_P^\star$), isothermal compressibility ($\kappa_T^\star$) and isobaric heat capacity ($c_P^\star$) by computing the covariance and variance in the volume and enthalpy (H) as

$$\alpha_P = \frac{\langle HV \rangle - \langle H \rangle \langle V \rangle}{N \langle V \rangle k_B T^2} = \frac{k_B}{\varepsilon_{BB}} \alpha_P^\star, \quad (10)$$

$$\kappa_T = \frac{\langle V^2 \rangle - \langle V \rangle^2}{N \langle V \rangle k_B T} = \frac{\sigma^3}{\varepsilon_{BB}} \kappa_T^\star, \quad (11)$$

$$c_P = \frac{\langle H^2 \rangle - \langle H \rangle^2}{N k_B T^2} = k_B c_P^\star. \quad (12)$$

These thermodynamic quantities, as well as the average density, were computed from five independent NPT simulations, consisting of 1×10^8 Monte Carlo cycles, following equilibration at each temperature and pressure considered.

Molecular dynamics simulations of TIP4P/Ice and TIP4P/2005. Molecular dynamics simulations were performed for the TIP4P/Ice[53] and TIP4P/2005[54] water models in the NPT ensemble, using GROMACS.[62] The integration time step was 2 fs. The Nosé–Hoover thermostat and the isotropic Parrinello–Rahman barostat, both with characteristic times of a few picoseconds, were used. Rigid constraints for the molecular model were implemented with a sixth-order linear constraint solver. The long-range electrostatic interactions were dealt with by using the particle-mesh Ewald method at fourth order. The cut-off distances for both the van der Waals and the real-space electrostatic interactions were fixed at 0.9 nm. For the TIP4P/Ice model, we focused on the $T = 188$ K isotherm for systems of $N = 300$ and 1,000 molecules, with pressures ranging from 1 bar to 4,000 bar. For the TIP4P/2005 model, we performed simulations at $T = 177$ K and $P = 1{,}750$ bar for a system of $N = 300$ molecules. The analysed trajectories were longer than 40 µs for both models, and equilibrated data from ref. [9] were used as starting configurations.

Cluster identification in the colloidal water model. We identified the tetrahedral clusters formed by the triblock patchy particles using the local order parameter, q_{Td} (which equals 1 for a perfect tetrahedron), defined as[63]

$$q_{Td} = 1 - \frac{3}{8} \sum_{i=1}^{N_c-1} \sum_{j=i+1}^{N_c} \left(\cos\psi_{ij} + \frac{1}{3}\right)^2, \quad (13)$$

where $N_c = 4$ is the number of particles in the cluster under consideration and ψ_{ij} is the angle subtended at the centre of the cluster by the two vectors joining the centre of particle i and particle j. We defined the set of neighbours for each particle i as those with which it shares a patch A–patch A bond. q_{Td} was then calculated for all distinct combinations of four particles including i from this set. All unique sets of four particles satisfying the condition $q_{Td} > 0.99$ were considered to form a tetrahedron.

Multiple histogram reweighting. To identify the location of the LLCP in the colloidal water system, we merged information collected from the Monte Carlo simulations to generate an estimate of the density of states, $\hat{\Omega}(E, V)$. In the NPT ensemble, the partition function at a given pressure P and temperature T can be written as[9,64]

$$\mathcal{Z}(P, \beta) = \sum_E \sum_V \hat{\Omega}(E, V) e^{-\beta(E+PV)} = e^{-\beta G(P,\beta)}, \quad (14)$$

where E is the energy of the system, V is the volume and G is the Gibbs free energy. We estimated the density of states by performing R simulations at different temperatures and pressures, where the total set of state points considered was given by $\{T_i, P_i\}$, $i = 1, \ldots, R$. For each simulation, a histogram (H_i) of the energies and volumes sampled was generated using data from C_i configurations spaced equally along the trajectory. We defined an estimate for the density of states at discrete values of energy and volume, using data from a single simulation, as

$$\Omega_i(E, V) = \frac{H_i(E, V)}{C_i} e^{+\beta_i(E + P_i V)} e^{-\beta_i G(P_i, \beta_i)}. \tag{15}$$

However, each of the R simulations only provides information of the density of states for a finite portion of the phase space. Hence, a better estimate was obtained by combining all the information gathered from the R simulations. This 'best guess' for the density of states was written as a linear combination of these individual estimates, combined with a weight factor $\omega_i(E, V)$, as

$$\hat{\Omega}^b(E, V) = \sum_{i=1}^{R} \omega_i(E, V) \Omega_i(E, V)$$
$$= \sum_{i=1}^{R} \omega_i(E, V) \frac{H_i(E,V)}{C_i} e^{+\beta_i(E+P_iV)} e^{-\beta_i G(P_i, \beta_i)} \tag{16}$$

Conventionally, it is assumed that the 'random' variable $H_i(E, V)$ has Poisson statistics. It then becomes possible to optimize equation (16) by minimizing the variance of $\hat{\Omega}^b(E, V)$ using the method of Lagrange multipliers, under the constraint that $\sum_i^R \omega_i = 1$. As a result

$$\hat{\Omega}^b(E, V) = \frac{\sum_{i=1}^{R} H_i(E, V)}{\sum_{i=1}^{R} e^{-\beta_i(E + P_i V)} e^{+\beta_i G(P_i, \beta_i)} C_i} \tag{17}$$

and

$$e^{-\beta_i G(P_i, \beta_i)} = \sum_E \sum_V \hat{\Omega}^b(E, V) e^{-\beta_i(E + P_i V)}. \tag{18}$$

When the number of samples used to generate each of the R histograms is the same (that is, $\forall i: C_i = C$), the above equations can be simplified by writing

$$S(E, V)/k_B = \ln \hat{\Omega}^b(E, V) = \ln \left(\frac{\sum_{i=1}^{R} H_i(E,V)}{C} \right)$$
$$- \ln \left(\sum_{i=1}^{R} e^{-\beta_i(E + P_i V)} e^{+\beta_i G(P_i, \beta_i)} \right) \tag{19}$$

and

$$G(P_i, \beta_i) = -k_B T \ln \left(\sum_E \sum_V e^{S(E,V)/k_B - \beta_i(E + P_i V)} \right). \tag{20}$$

Evaluation of the exponential terms can be numerically unstable, but rewriting these equations in terms of the entropy allows us to make use of $\ln(e^x) = a + \ln(e^{x-a})$ to improve the numerical precision. Equations (19) and (20) can then be solved self-consistently via an iterative procedure, taking $G_i = 0$ for all i as an initial guess. We assume that the values of the Gibbs free energies have converged once the sum squared difference between consecutive estimates is less than a tolerance value of 10^{-7}. The converged solutions provided an estimate for the values of $G(P, \beta)$ up to some additive constant c. From these estimates, we obtained values for the density of states. It was, therefore, possible to predict the probability of observing a state with energy E and volume V at any relevant state point (P, β) using

$$e^{-\beta G(P, \beta) + c} = \sum_E \sum_V e^{S(E,V)/k_B - \beta(E + PV)}$$
$$\Rightarrow p(E, V | \beta, P) = \frac{e^{\mathcal{c}} e^{S(E,V)/k_B - \beta(E + PV)}}{\sum_E \sum_V e^{\mathcal{c}} e^{S(E,V)/k_B - \beta(E + PV)}} \tag{21}$$

Bond definition. In the case of colloidal water, the network is created by the patch B–patch B interactions, forming bonds that join discrete tetrahedral clusters. Owing to the square-well nature of the patch–patch interaction, searching for the existing bonds becomes equivalent to searching for all pairs of particles with pair interaction energy equal to $-\epsilon_{BB}$.

For the TIP4P/Ice and TIP4P/2005 models, a bond between two water molecules was deemed to exist when a hydrogen bond was present. This was defined on the basis of a geometric criterion[65] by the conditions $r < 0.35$ nm and $\theta < 30°$. Here, r is the intermolecular oxygen–oxygen distance and θ is the smallest among the four angles between the intramolecular O–H and the intermolecular O–O lines. As shown previously[19], no ambiguity in the identification of hydrogen bonds remains if the Chandler–Luzar geometric criterion is applied using inherent structure configurations. These are the local potential energy minimum

configurations reached via a steepest descent path that quenches the vibrational degrees of freedom. To evaluate the inherent structures, we used the steepest descent algorithm in GROMACS (compiled in double precision) with a force tolerance of $1 \, \text{J mol}^{-1} \, \text{nm}^{-1}$ and a maximum step size of 5×10^{-4} nm.

Ring statistics. We calculated the number of rings ($\mathcal{N}_{R_l}$) with sizes $l \in [4 - l_{max}]$ where each bond network was abstracted into a periodic undirected graph G. The vertices (VG) of the graph represent the positions of the objects of interest (either the tetrahedral cluster centres or the oxygen atoms of the water molecules), and each of the edges (EG) connects two bonded objects as defined above. We defined a ring as a closed path in this graph (that is, a path whose first and final vertices are the same). We excluded rings in which non-adjacent vertices are directly connected via an edge. For each vertex V_G^i in the graph, we determined all distinct rings, sequentially identifying all rings associated with each vertex, using a depth-first search traversal, up to the chosen l_{max}. To improve the efficiency of the analysis, following the extraction of all relevant paths starting from V_G^i (and hence rings containing V_G^i), all edges connected to that vertex were removed from the graph before initiating the depth-first search traversal from V_G^{i+1}. We also accounted for the double counting of a ring, which arises from the undirected nature of the graph.

Linking number and writhe calculations. We measured the entanglement of liquid networks using metrics derived from the helicity, which is defined in equation (2), in terms of Lk_{ij}, the linking number between two independent closed curves R_i and R_j, and Wr_i, the writhe of a single closed curve. Both the linking number and the writhe can be defined as double line integrals[46,66–68]:

$$Lk_{ij}(R_i, R_j) = \frac{1}{4\pi} \oint_{R_i} \oint_{R_j} \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3} \cdot (d\mathbf{r}_j \times d\mathbf{r}_i), \tag{22}$$

$$Wr_i(R_i) = \frac{1}{4\pi} \oint_{R_i} \oint_{R_i} \frac{\mathbf{r}_i' - \mathbf{r}_i}{|\mathbf{r}_i' - \mathbf{r}_i|^3} \cdot (d\mathbf{r}_i' \times d\mathbf{r}_i). \tag{23}$$

Note that equation (22) and (23) are combined together in equation (1) (the contributions $i \neq j$ and $i = j$, respectively). The closed curves from which we computed $\mathcal{L}_n$ and $\mathcal{W}_n$ are constructed from the set of rings identified in the liquid networks from both models, where $\mathbf{r}_i$ and $\mathbf{r}_j$ represent the positions of the objects that make up rings R_i and R_j, respectively. We calculated the linking number only for pairs of disjoint rings (rings not sharing any vertices), and the writhe for both knotted individual rings and for pairs of rings that share at least one edge that were also knotted (these pairs of rings can essentially be considered as composite rings with intra-ring edges).

Computing the linking number. The rings are composed of discrete line segments and so can be considered as closed polygonal paths. As a result, the linking number can be computed using the expression for polylines[69]:

$$Lk_{ij}(R_i, R_j) = \frac{1}{4\pi} \sum_{kl} \lambda_{kl}, \tag{24}$$

where λ_{kl} is the contribution to the Lk_{ij} from a pair of line segments k and l belonging to rings i and j, respectively, and

$$\lambda_{kl} = \text{atan} \left(\frac{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}{|\mathbf{a}||\mathbf{b}||\mathbf{c}| + (\mathbf{a} \cdot \mathbf{b})|\mathbf{c}| + (\mathbf{c} \cdot \mathbf{a})|\mathbf{b}| + (\mathbf{b} \cdot \mathbf{c})|\mathbf{a}|} \right)$$
$$+ \text{atan} \left(\frac{\mathbf{c} \cdot (\mathbf{d} \times \mathbf{a})}{|\mathbf{c}||\mathbf{d}||\mathbf{a}| + (\mathbf{c} \cdot \mathbf{d})|\mathbf{a}| + (\mathbf{a} \cdot \mathbf{c})|\mathbf{d}| + (\mathbf{d} \cdot \mathbf{a})|\mathbf{c}|} \right), \tag{25}$$

where $\mathbf{a} = \mathbf{r}_l - \mathbf{r}_k$, $\mathbf{b} = \mathbf{r}_l - \mathbf{r}_{k+1}$, $\mathbf{c} = \mathbf{r}_{l+1} - \mathbf{r}_{k+1}$ and $\mathbf{d} = \mathbf{r}_{l+1} - \mathbf{r}_k$ (ref. [69]).

Computing the writhe. The writhe has previously been used to study folding in protein systems[70] and supercoiling in DNA[68]. In these studies, proteins and DNA strands are modelled as polymer chains made up of straight line segments. Similarly, we computed the writhe using[68]

$$Wr_i(R_i) = \frac{1}{2\pi} \sum_{k \neq l} \omega_{kl}, \tag{26}$$

where ω_{kl} is the contribution to the writhe from the segments k and l on ring i, defined as[68]

$$\omega_{kl} = \text{sgn} \left\{ [(\mathbf{r}_{l+1} - \mathbf{r}_l) \times (\mathbf{r}_{k+1} - \mathbf{r}_k)] \cdot \mathbf{a} \right\} \left(\sin^{-1}(\mathbf{n}_1 \cdot \mathbf{n}_2) + \sin^{-1}(\mathbf{n}_2 \cdot \mathbf{n}_3) \right.$$
$$\left. + \sin^{-1}(\mathbf{n}_3 \cdot \mathbf{n}_4) + \sin^{-1}(\mathbf{n}_4 \cdot \mathbf{n}_1) \right), \tag{27}$$

where $\text{sgn}(\cdot)$ is the sign function and

$$\mathbf{n}_1 = \frac{\mathbf{a} \times \mathbf{d}}{|\mathbf{a} \times \mathbf{d}|}, \quad \mathbf{n}_2 = \frac{\mathbf{d} \times \mathbf{c}}{|\mathbf{d} \times \mathbf{c}|}, \quad \mathbf{n}_3 = \frac{\mathbf{c} \times \mathbf{b}}{|\mathbf{c} \times \mathbf{b}|}, \quad \mathbf{n}_4 = \frac{\mathbf{b} \times \mathbf{a}}{|\mathbf{b} \times \mathbf{a}|}, \tag{28}$$

where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ are defined as above.

Identifying knots. We made use of the freely available Python package 'pyknotid' to identify whether a ring (or a pair of rings sharing vertices) is knotted[71]. Using 'pyknotid', we computed the Gauss code for the ring, and following simplification by performing Reidemeister moves, we identified whether a knot was present. If present, the identity of the knot was determined using a knot look-up table, which includes information about all knots with up to 15 crossings. Only if the ring (or a pair of rings sharing vertices) was deemed to be knotted was the writhe then calculated.

Data availability

Datasets used to generate the figures in this article have been deposited in the University of Birmingham edata Repository and can be accessed from https://edata.bham.ac.uk/828/.

Code availability

Our Monte Carlo code, developed in-house, can be accessed at https://github.com/cdwaipayan/PaSSion.git.

References

58. Bol, W. Monte Carlo simulations of fluid systems of waterlike molecules. *Mol. Phys.* **45**, 605–616 (1982).
59. Kern, N. & Frenkel, D. Fluid–fluid coexistence in colloidal systems with short-ranged strongly directional attraction. *J. Chem. Phys.* **118**, 9882–9889 (2003).
60. Frenkel, D. & Smit, B. *Understanding Molecular Simulation: from Algorithms to Applications*, vol. 1, 2nd edn (Academic Press, 2001).
61. Almarza, N. G. A cluster algorithm for Monte Carlo simulation at constant pressure. *J. Chem. Phys.* **130**, 184106 (2009).
62. Abraham, M. J. et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
63. Errington, J. R. & Debenedetti, P. G. Relationship between structural order and the anomalies of liquid water. *Nature* **409**, 318–321 (2001).
64. Bianco, V. & Franzese, G. Critical behavior of a water monolayer under hydrophobic confinement. *Sci. Rep.* **4**, 1–10 (2014).
65. Luzar, A. & Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **379**, 55–57 (1996).
66. Gauss, C. F. & zu Göttingen, K. G. d. W. Zur mathematischen theorie der electrodynamischen wirkungen. In Werke, 601–630 (Springer, 1877).
67. Fuller, F. B. The writhing number of a space curve. *Proc. Natl Acad. Sci. U. S. A.* **68**, 815–819 (1971).
68. Klenin, K. & Langowski, J. Computation of writhe in modeling of supercoiled DNA. *Biopolym. Orig. Res. Biomol.* **54**, 307–317 (2000).
69. Qu, A. & James, D. L. Fast linking numbers for topology verification of loopy structures. *ACM Trans. Graph.* **40**, 106 (2021).
70. Levitt, M. Protein folding by restrained energy minimization and molecular dynamics. *J. Mol. Biol.* **170**, 723–764 (1983).
71. Taylor, A. J. & other SPOCK contributors. pyknotid knot identification toolkit. *GitHub* https://github.com/SPOCKnots/pyknotid (2017).

Author contributions

A.N., D.C. and F.S. designed the research. A.N. and F.S. performed the research. A.N. contributed new analysis tools. A.N., D.C. and F.S. analysed the data. A.N., D.C. and F.S. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41567-022-01698-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-022-01698-6.

Correspondence and requests for materials should be addressed to Andreas Neophytou, Dwaipayan Chakrabarti or Francesco Sciortino.

Peer review information *Nature Physics* thanks Camille Scalliet and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
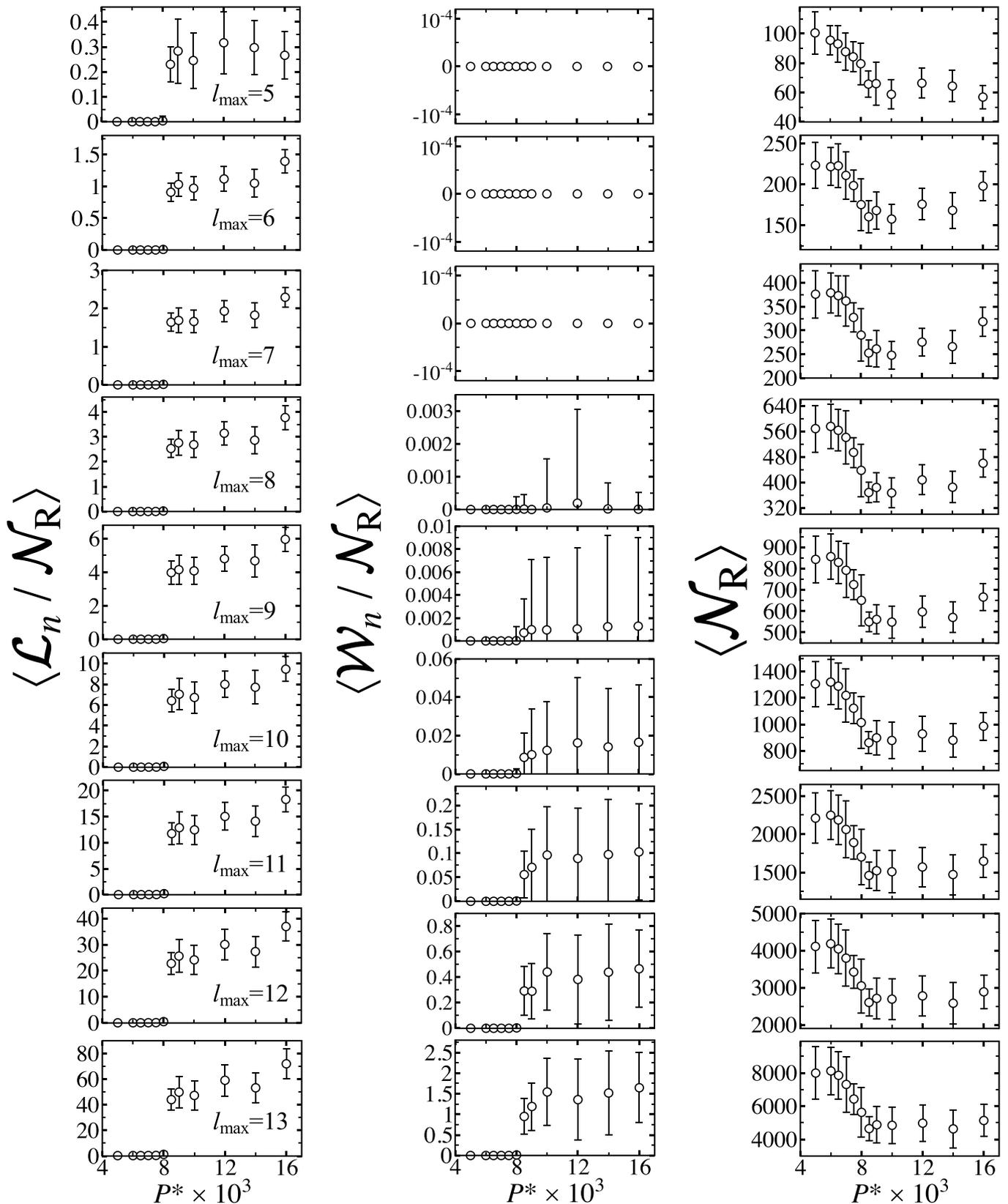
Extended Data Fig. 1 | Pressure dependence of the topological properties of the colloidal water model. Pressure dependence of the topological properties of the colloidal water model. Pressure dependence of $\langle \mathcal{L}_n/\mathcal{N}_R \rangle$, $\langle \mathcal{W}_n/\mathcal{N}_R \rangle$ and $\langle \mathcal{N}_R \rangle$ for the colloidal water model, for $5 \leqslant l_{max} \leqslant 13$ at a temperature of $T^\star = 0.105$. Note that $\mathcal{W}_n = 0$ for all pressures when $l_{max} \leqslant 7$.
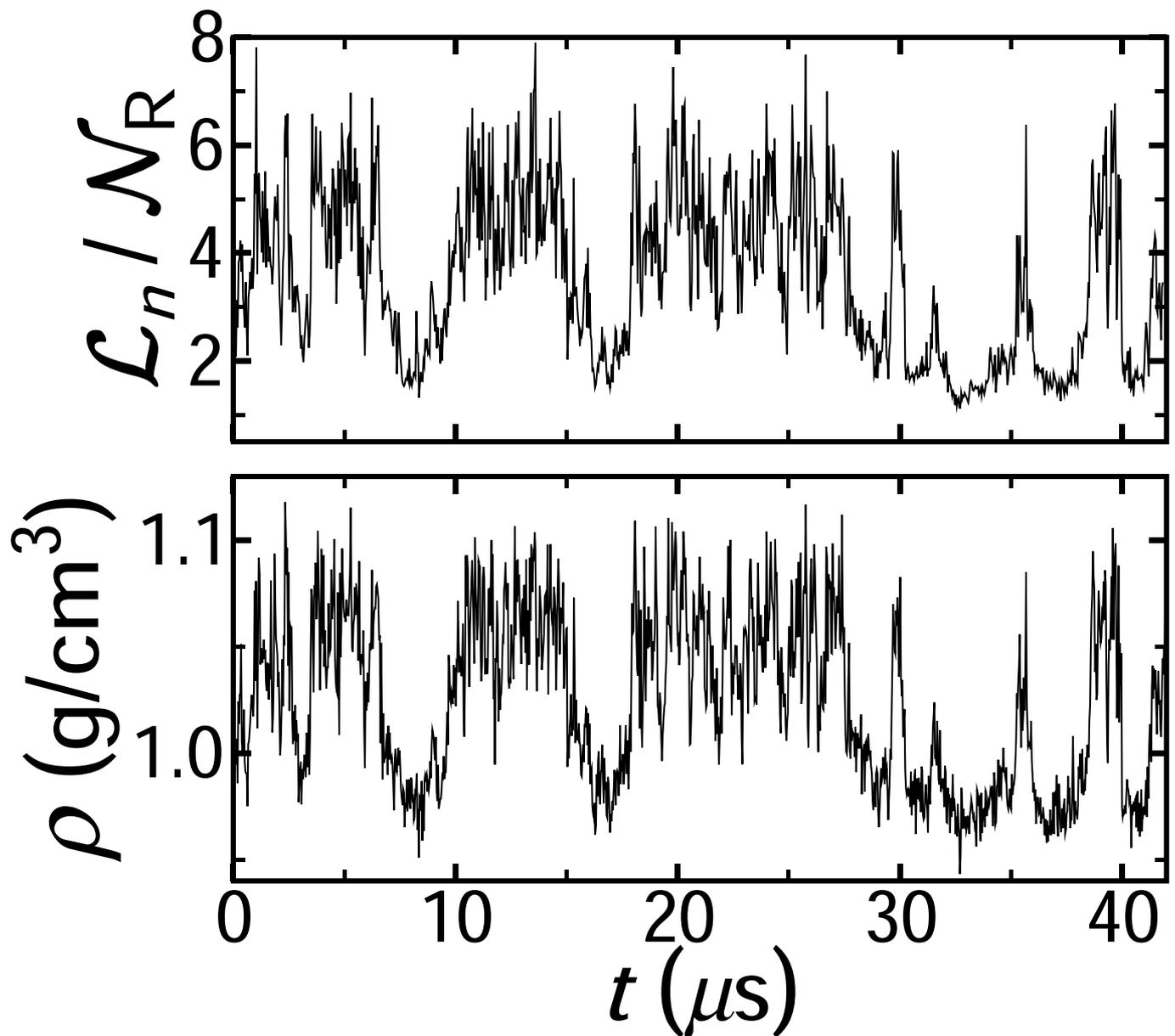
Extended Data Fig. 2 | Topological nature of the liquid-liquid phase transition in TIP4P/2005. Topological nature of the liquid-liquid phase transition in TIP4P/2005. Fluctuations in density (ρ) and $\mathcal{L}_n$ (computed using $l_{max=13}$) with time (t) along an isobaric-isothermal molecular dynamics trajectory for $N=300$ TIP4P/2005 water molecules at a temperature of $T=177$K and a pressure of $P=1750$bar, a state point close to its liquid-liquid critical point.