

# Structural Order as the Key Phase Indicator in Supercooled Liquid Water

Saeed Pourasad<sup>1</sup> and Kwang S. Kim<sup>1</sup>✉

## 1 Materials and Methods

### 1.1 Simulation Model

We performed simulations using the TIP4P/2005 water model,<sup>1</sup> a re-parametrized version of the ‘classical’ TIP4P model. Widely utilized, this model has demonstrated exceptional reliability and accuracy in capturing various experimental characteristics of water.

### 1.2 Simulation Methods

We conducted molecular dynamics simulations of water using the LAMMPS<sup>2</sup> package, employing both the isothermal-isochoric (*NVT*) and the isothermal-isobaric (*NPT*) ensembles. In both cases, we utilized the Verlet integrator with a time step of 1 fs (and 0.1 fs for some cases). Electrostatic interactions were computed using the Particle-Particle Particle-Mesh (PPPM) method with a real space cutoff distance of 1.2 nm. Similarly, a 1.2 nm cutoff distance was applied for van der Waals interactions. For both *NVT* and *NPT* simulations, we employed the Nose–Hoover thermostat and barostat. Cubic systems (with periodic boundary conditions) of size 1024 and cubic systems (with periodic boundary conditions) of sizes 512 and 1620 were studied in *NVT* and *NPT* Simulations, respectively.

### 1.3 Generalized Replica Exchange Method

The generalized replica exchange method (gREM) is tailored for simulating systems undergoing first-order phase transitions. In a *NVT* simulation, the probability distribution of the potential energy,  $P(E)$  is defined as  $P(E) = W(E) \Omega(E)/Z$ , where  $W(E) \equiv \exp(-w(E))$  is the sampling weight,  $\Omega(E)$  is the density of states (linked to entropy,  $S(E)$ ), through  $\Omega(U) = \exp(-S(E)/k_B)$ ), and  $Z = \int dU W(E) \Omega(E)$  represents the generalized configurational integral.

The generalized ensemble (g) introduces a generalized temperature function,  $dw(E)/dE = 1/(k_B T^g(E))$  with  $w = \beta E$ ,  $\beta = 1/(k_B T)$ , and  $T^g$  is the usual constant temperature  $T$ . Computing the extrema,  $E^*$  of  $P(E)$  yields  $T^g(E^*) = T_S(E^*)$ , where  $T_S(E) = 1/(dS(E)/dE)$  is the statistical temperature. At a first-order phase transition, the statistical temperature exhibits an *S*-shaped loop, attributed to the “convex intruder” in entropy stemming from the reduced density of states in the phase coexistence range due to surface effects.

In the vicinity of the transition region, within a constant temperature ensemble,  $T_S(E^*)$  has three solutions corresponding to two pure phases at  $E_1$  and  $E_2$ , and the barrier at  $E_{bar}$ . These points align with two minima and

---

<sup>1</sup>Center for Superfunctional Materials, Department of Chemistry, Ulsan National Institute of Technology and Science (UNIST), Ulsan 44919, Republic of Korea. ✉ email: kimks@unist.ac.kr

one maximum in the Gibbs free-energy profile at the specified temperature. States around the barrier are seldom visited, with the system predominantly residing in one of the free-energy minima ( $E_1$  or  $E_2$ ).

The gREM conceptually involves employing an ensemble with a temperature function that intersects  $T_S$  only once. This approach facilitates the sequential exploration of the unstable energy region between two stable states. This generalized temperature function as the linear effective temperature is defined as follows:

$$T_\alpha(E) = \lambda_\alpha + \gamma_0(E - E_0) \quad (1)$$

where the slope,  $\gamma_0$ , should be sufficiently negative to ensure a single intersection with  $T_S$  despite the  $S$  loop. By varying the intercept  $\lambda_\alpha$ , the intersection point with  $T_S$  can be systematically adjusted through the coexistence range, enabling thorough sampling with a series of unimodal, overlapping energy distributions. Replica exchanges are periodically attempted, with acceptance probabilities chosen to preserve detailed balance. Thermodynamic quantities can be computed using the statistical temperature, accessible from energy histograms in the replicas through the statistical temperature-weighted histogram analysis method (ST-WHAM).

In this particular case, 16 generalized replicas (ensembles) of water molecules ( $N=1024$ ) at a density of  $\rho = 1.0 \text{ g/cm}^3$  were utilized. A swap between adjacent ensembles was attempted every 1000 time steps. Each production run spanned 60-80 ns, during which the potentials of the configurations were sampled every 10 time steps for subsequent analysis.

In the gREM, a walker is an independent simulation trajectory that explores the phase space of the system. A replica, on the other hand, represents a version of the system defined by particular parameters, such as temperature, pressure, or Hamiltonian settings. Walkers are assigned to replicas during the simulation, and exchanges between replicas allow walkers to transition between different conditions, allowing better sampling of the system's landscape and overcoming energy barriers. This approach ensures efficient exploration while maintaining the desired ensemble distribution.

## 1.4 Well-Tempered Metadynamics

Well-tempered metadynamics (WTMetaD)<sup>3</sup> revolutionizes molecular dynamics simulations by systematically filling energy minima encountered during the simulation with a bias potential. This process effectively smooths out the free energy surface, facilitating the sampling of rare events crucial for understanding molecular behavior. Unlike conventional metadynamics, WTMetaD introduces a pivotal component known as the "well-tempering" parameter. This parameter regulates the rate at which the bias potential is added, ensuring a gradual and controlled convergence over time. By doing so, WTMetaD prevents the excessive oversampling of certain states, promoting a more efficient exploration of the system's configurational space. To implement this method, the PLUMED package<sup>4</sup> is integrated with the LAMMPS package.

WTMetaD achieves this by incorporating Gaussian-shaped bias potentials  $W(\phi, t)$ , strategically centered around previously explored regions of the energy landscape. As the simulation progresses, these Gaussians accumulate, flattening the energy landscape and enabling the system to explore a broader range of states. Crucially, the "bias factor" within WTMetaD plays a fundamental role in this process. It governs the rate at which these Gaussian bias potentials are introduced, effectively influencing the pace of exploration and ensuring a balanced sampling across different states. By finely adjusting the bias factor, we can maintain a uniform exploration of the energy landscape, avoiding the overemphasis on specific regions while promoting an efficient and comprehensive analysis of molecular dynamics. For any time  $t$ , the cumulative bias potential is computing using:

$$W(\phi, t) = W_0 \sum_{t' \leq t} \exp\left[-\frac{W(\phi, t')}{k_B \Delta T}\right] \exp\left(-\sum_{i=1}^n \frac{[\phi_i - \phi_i(t')]^2}{2\sigma\phi_i^2}\right) \quad (2)$$

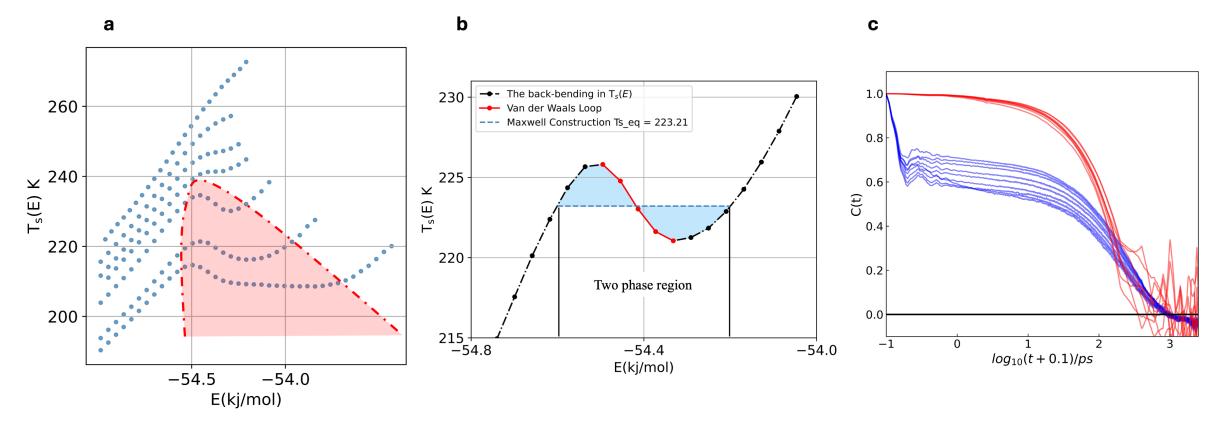
where  $n$  is the number of order parameters, parameters  $\sigma\phi_i^2$  and  $W_0$  set the width and initial height of the Gaussians respectively, and parameter  $\Delta T$ , as we mentioned earlier controls the rate of decay of their height. Values for the parameters in eqn (3) were systematically determined using the procedure developed by Quigley.<sup>5</sup> This resulted in  $W_0 = 0.2 k_B T$  and the Gaussian width that is estimated from the fluctuations of the collective variable (volume) in an unbiased simulation. Additionally, we updated the bias factor in our study, increasing it from 6 to 11.

## 1.5 Bootstrap Analysis

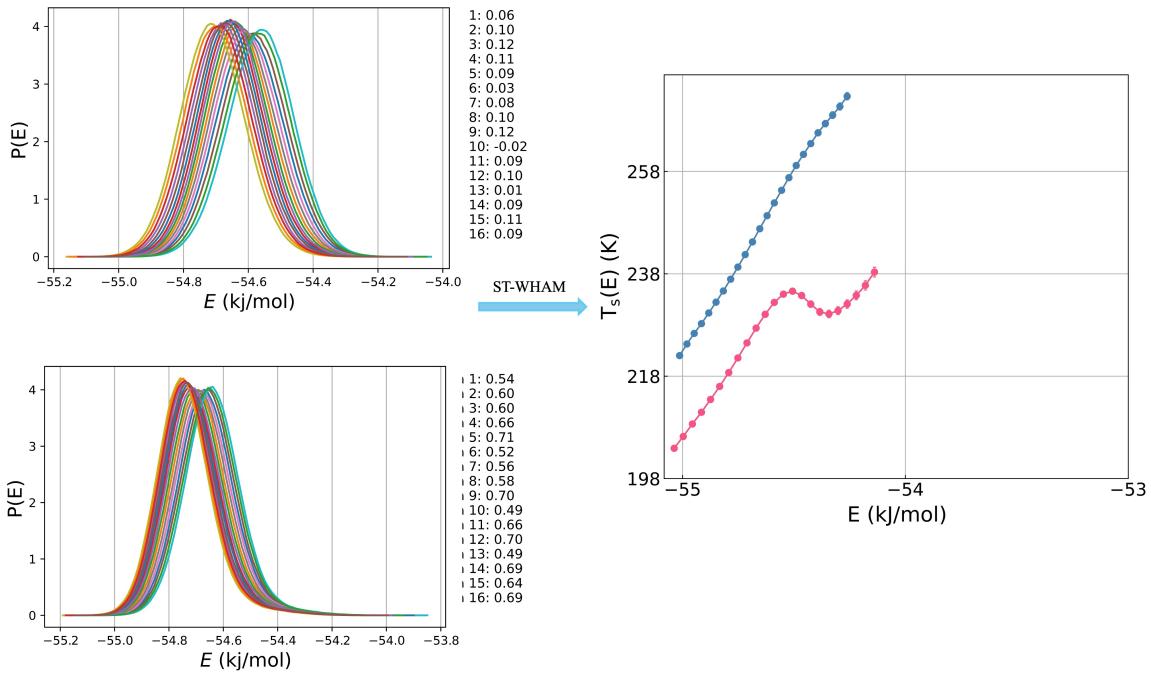
Bootstrap analysis<sup>6</sup> is a statistical method utilized to estimate the sampling distribution and ascertain its level of uncertainty. This technique generates numerous resamples by randomly selecting observations from the initial sample, with replacement. Consequently, an observation may be chosen multiple times within the same bootstrap sample, while others may not be selected at all. Each bootstrap sample maintains the same size as the original sample, meaning that if there are " $n$ " observations in the original data set, each bootstrap sample also encompasses " $n$ " observations. We then calculate the average of these re-sampled data sets, repeat this process numerous times, and subsequently determine the standard deviation of the averages obtained. This provides us with an estimate of the statistical uncertainty associated with the average computed using actual data.

## 1.6 Radial Distribution Function

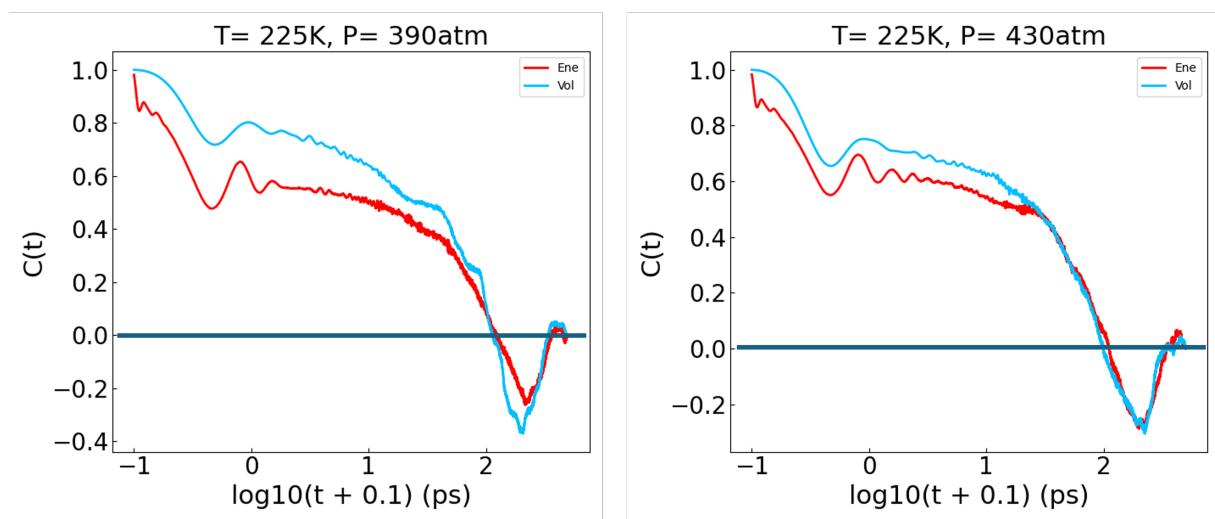
The structural distinctions among the two states are illustrated in the Extended Data Fig.4 through the radial distribution function.



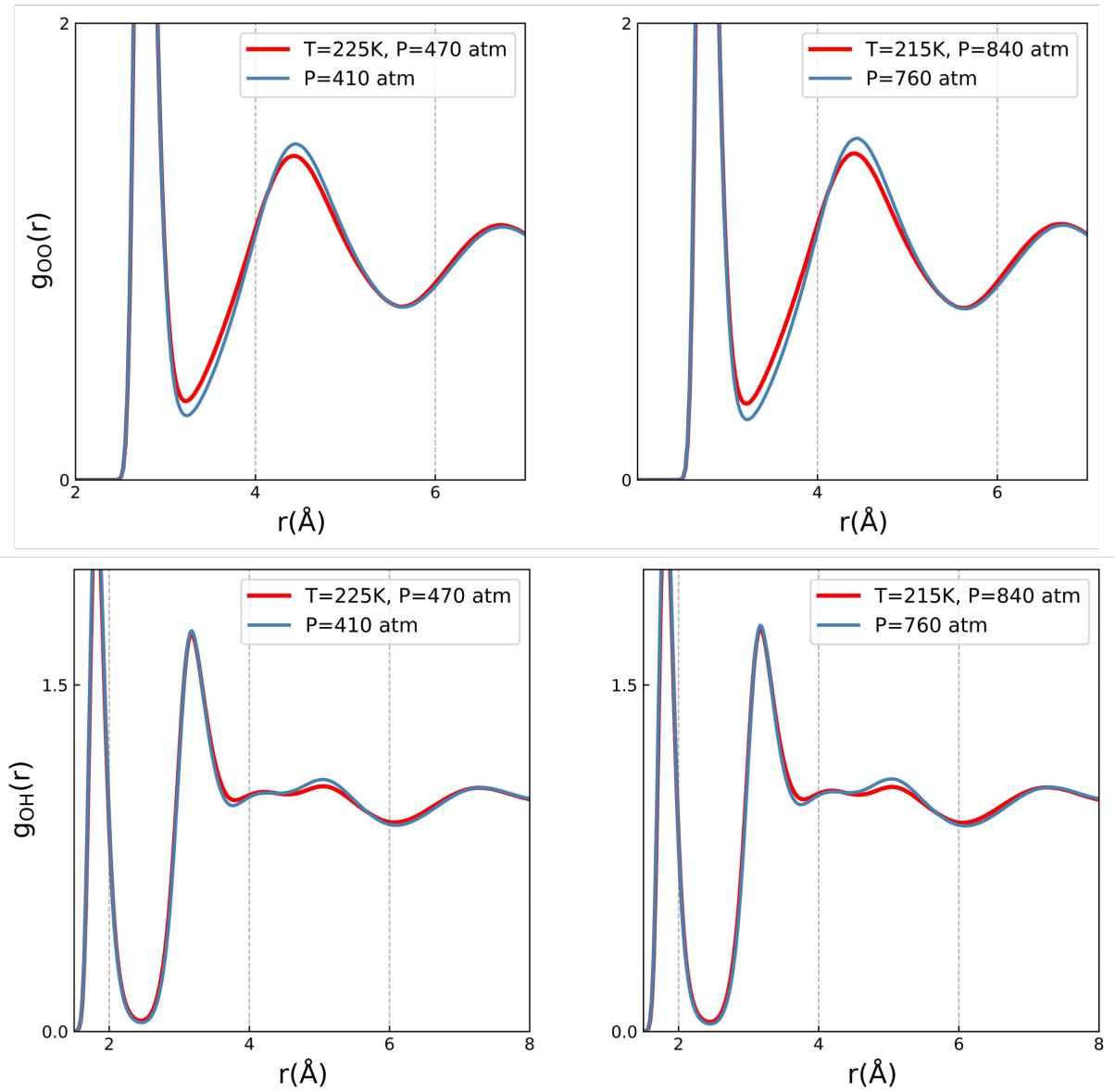
**Supplementary Fig. 1** | a) The statistical temperature-total energy graphs were computed using the gREM method. The two-phase region is schematically highlighted in red, extending up to the critical point. b) The essential conditions below the actual critical point are illustrated. c) The autocorrelation functions of various walkers (red) and replicas (blue) (from graph b) are depicted.



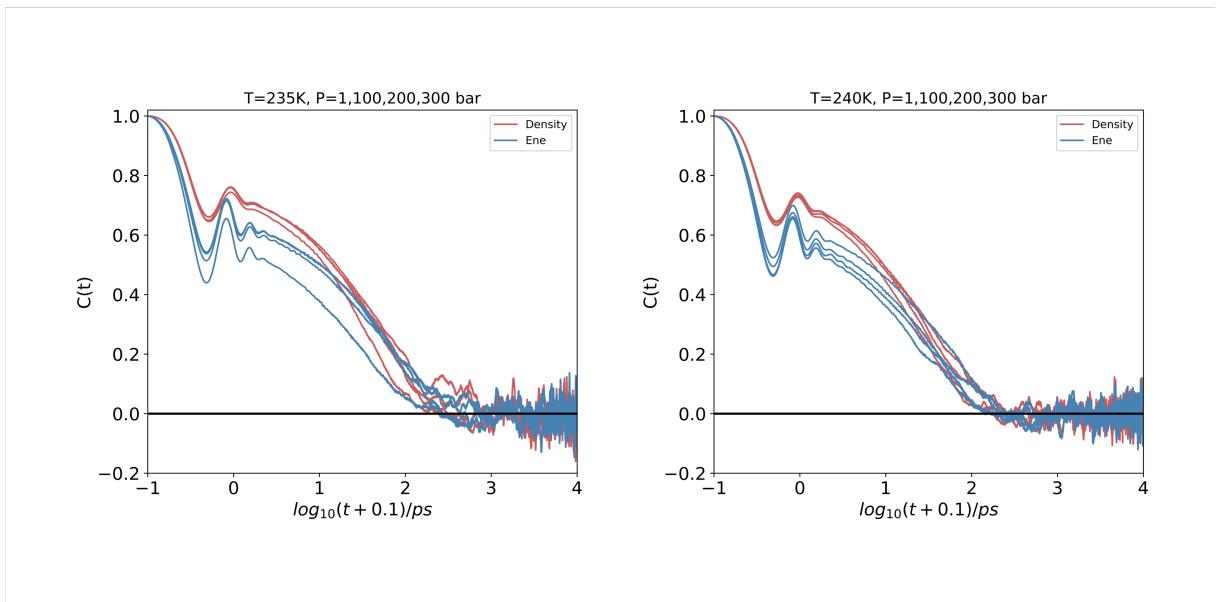
**Supplementary Fig. 2** | Total energy histograms of 16 replicas with different temperature ranges  $T$ ,  $T+2, \dots, T+30$ K. The uppermost graph corresponds to the higher temperature. The skewness of each replica at each graph is presented in the column next to the graphs.



**Supplementary Fig. 3** | The autocorrelation function of unbiased NPT simulations at constant temperature and two different pressures.



**Supplementary Fig. 4** | Radial distribution functions for the highly ordered and less ordered liquid states, extracted from unbiased NPT simulations at two different temperatures.



**Supplementary Fig. 5** | The autocorrelation function from unbiased NPT simulations at two different temperatures and various pressures near the inflection point of the gREM.

## References

- [1] Abascal, J. L. F. & Vega, C. A general purpose model for the condensed phases of water: Tip4p/2005. *J. Chem. Phys.* **123**, 234505 (2005).
- [2] Thompson, A. P. *et al.* Lammps - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **271**, 108171 (2022).
- [3] Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603 (2008).
- [4] Bonomi, M. *et al.* Plumed: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **180**, 1961–1972 (2009).
- [5] Quigley, D. & Rodger, P. A metadynamics-based approach to sampling crystallisation events. *Molecular Simulation* **35**, 613–623 (2009).
- [6] Hub, J. S., de Groot, B. L. & van der Spoel, D. g\_wham—a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *Journal of Chemical Theory and Computation* **6**, 3713–3720 (2010).