

# Solvent Structure around Nanoparticles

## Investigating Hydrophobicity via MD Simulations

Shuvam Banerji Seal (22MS076)

Indian Institute of Science Education and Research Kolkata

November 19, 2025

Repository: [git@github.com:Shuvam-Banerji-Seal/Solvent-Structure-around-NanoParticles.git](https://github.com/Shuvam-Banerji-Seal/Solvent-Structure-around-NanoParticles.git)

## Table of Contents I

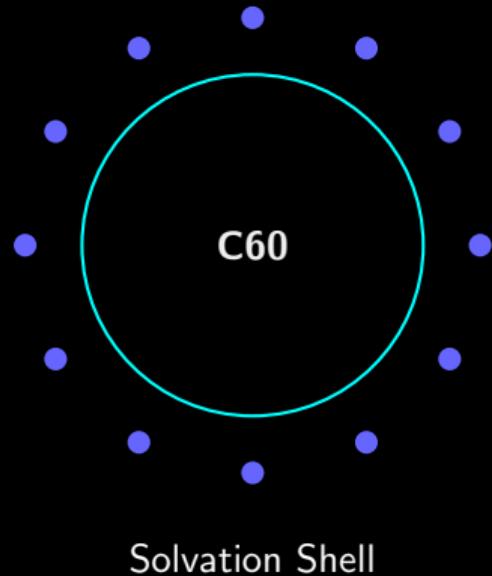
## Problem Statement

### Core Scientific Question:

- ▶ How does the solvent (water) reorganize itself around a nanoparticle ( $C_{60}$ )?
- ▶ How does this reorganization change as we tune the interaction strength ( $\epsilon_{CO}$ )?
- ▶ *Hypothesis:* The transition from hydrophobic to hydrophilic induces a structural phase transition in the solvation shell.

### The Machine Learning Angle:

- ▶ Can we generate descriptors (RDF fingerprints, Voronoi tessellation) from MD trajectories?
- ▶ Can ML models predict aggregation propensity based purely on local solvent structure?



Solvation Shell

### Composition:

- ▶ **Solute:** 3 × Buckminsterfullerene (C<sub>60</sub>) Molecules ( $N_{atoms} = 180$ ).
- ▶ **Solvent:** ~2000 Water molecules (TIP4P/2005 model).
- ▶ **Box:** Cubic 40 × 40 × 40 Å<sup>3</sup> (Periodic Boundary Conditions).

### Physics Engine (LAMMPS):

- ▶ **Force Field:** Lennard-Jones + Coulombic (Long Range).
- ▶ **Thermostat/Barostat:** Nosé-Hoover (NVT/NPT).
- ▶ **Constraints:** SHAKE algorithm for rigid water bonds.

## Step 1: Raw Data (`C60.data`)

**Origin:** Derived from the standard XYZ coordinates available in the LAMMPS library.

### **Content:**

- ▶ Contains spatial coordinates ( $x, y, z$ ) for 60 Carbon atoms.
- ▶ *Limitation:* Purely geometric. No bonding topology (bonds, angles, dihedrals) is defined in the raw file.

*We must computationally infer the topology to treat C<sub>60</sub> as a rigid, bonded molecule.*

## Step 2: Topology Generation (generate\_C60\_bonded.py)

**Objective:** Transform point cloud to bonded graph. **Bond Criteria:** C-C bond length in Fullerene  $\approx 1.4 \text{ \AA}$ . Threshold:  $1.3 \leq r_{ij} \leq 1.5 \text{ \AA}$ .

```
1: Read Atoms  $A = \{a_1, \dots, a_{60}\}$ 
2:  $Bonds \leftarrow []$ 
3: for  $i \in 1 \dots 60$  do
4:   for  $j \in (i + 1) \dots 60$  do
5:      $d_{ij} = \sqrt{(x_i - x_j)^2 + \dots}$ 
6:     if  $1.3 \leq d_{ij} \leq 1.5$  then
7:        $Bonds.append((i, j))$ 
8:     end if
9:   end for
10: end for
11: Verify: Count bonds per atom
12: if  $\forall i, Count(i) == 3$  then
13:   Write C60_bonded.data
14: else
15:   Error: Invalid Topology
16: end if
```

**Mathematical Basis:** Euclidean Distance Metric:

$$r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|_2 \quad (1)$$

Graph Theory Constraint:

$$\deg(v) = 3, \quad \forall v \in V_{C60} \quad (2)$$

Total Bonds:

$$|E| = \frac{1}{2} \sum \deg(v) = \frac{60 \times 3}{2} = 90 \quad (3)$$

## Validation Script: Initialization & Force Field

### 1. Atom Definition:

```
1 |     units real          # Energy: kcal/mol, Dist: Angstrom, Time: fs
2 |     atom_style full      # Attributes: ID Mol Type Charge X Y Z
```

*Physics:* 'full' is required because TIP4P water relies on partial charges ( $q_O = -1.04e$ ,  $q_H = +0.52e$ ).

### 2. TIP4P/2005 Pair Style:

```
1 | pair_style lj/cut/tip4p/long/omp 2 3 2 1 0.1546 12.0
```

*Mathematical Model:* The potential includes a virtual site  $M$  located  $d_{OM} = 0.1546\text{\AA}$  from Oxygen along the H-O-H bisector.

$$\mathbf{r}_M = \mathbf{r}_O + d_{OM} \frac{\mathbf{r}_{H1} + \mathbf{r}_{H2} - 2\mathbf{r}_O}{\|\mathbf{r}_{H1} + \mathbf{r}_{H2} - 2\mathbf{r}_O\|} \quad (4)$$

The interaction energy is:

$$U_{ij} = 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_M q_j}{4\pi\epsilon_0 r_{Mj}} \quad (5)$$

### 3. Reading Topology:

```
1 |     read_data C60_bonded.data add merge  
2 |     molecule h2omol H2O_TIP4P2005_fixed.mol offset 1 1 0 0 0
```

*Logic:* Merges the C60 graph with the rigid water template.

### 4. Bonded Potentials (Harmonic Approximation):

```
1 |     bond_style harmonic ; bond_coeff 1 938.0 1.42  
2 |     angle_style harmonic ; angle_coeff 1 0.0 104.52
```

*Equation:* For C60 C-C bonds ( $k_b = 938 \text{ kcal/mol}/\text{\AA}^2$ ):

$$U_{bond}(r) = K_b(r - r_0)^2 \quad (6)$$

*Note:* Water bond coeffs are set to 0.0 because we use SHAKE (next slide).

## 5. Constraint Application:

```
1 | fix water_shake water_group shake 0.0001 20 0 b 2 a 1
```

**Mathematical Detail:** We treat water as a rigid body. Instead of solving Newton's equations directly, we solve the constrained Lagrangian:

$$\mathcal{L} = \sum \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - V(\mathbf{r}) - \sum_k \lambda_k \sigma_k(\mathbf{r}) \quad (7)$$

Where the holonomic constraints  $\sigma_k$  are:

$$\sigma_{OH} = \|\mathbf{r}_O - \mathbf{r}_H\|^2 - d_{OH}^2 = 0 \quad (8)$$

**Algorithm:** LAMMPS iteratively updates positions  $\mathbf{r}(t + \delta t)$  until the constraints are satisfied within tolerance ( $10^{-4}$ ).

### 6. Thermostatting:

```
1 |     velocity all create 300 12345 dist gaussian
2 |     fix 1 water_group nvt temp 300 300 100.0
```

**Nosé-Hoover Equations of Motion:** The code integrates these coupled differential equations:

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i} \quad (9)$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \zeta \mathbf{p}_i \quad (10)$$

$$\dot{\zeta} = \frac{1}{Q} \left( \sum \frac{\mathbf{p}_i^2}{m_i} - (3N + 1)k_B T_{target} \right) \quad (11)$$

where  $\zeta$  is the friction variable and  $Q$  is the thermal inertia parameter (controlled by 'Tdamp = 100.0').

### 7. Structural Integrity Check:

```
1 |     compute rg_c60_compute c60_group gyration  
2 |     fix rg_output all print 100 ...
```

**Metric Definition:** To ensure the C60 doesn't collapse or explode, we calculate the scalar Radius of Gyration  $R_g$ :

$$R_g = \sqrt{\frac{1}{M} \sum_{i=1}^{60} m_i (\mathbf{r}_i - \mathbf{r}_{COM}) \cdot (\mathbf{r}_i - \mathbf{r}_{COM})} \quad (12)$$

**Success Criteria:** For C60,  $R_g$  must remain strictly constant ( $\approx 3.5\text{\AA}$ ) if the bonds are defined correctly.

### 1. Box Definition:

```
1 |     region mybox block -20 20 -20 20 -20 20  
2 |     create_box 3 mybox ...
```

### Volume & Density Math:

- ▶ Box Side  $L = 40 \text{ \AA}$ .
- ▶ Volume  $V = L^3 = 64,000 \text{ \AA}^3$ .
- ▶ Target Density  $\rho \approx 1.0 \text{ g/cm}^3$ .
- ▶ Mass of Water  $M_w \approx 18 \text{ g/mol}$ .

Required number of water molecules  $N_w$ :

$$N_w = \frac{\rho V N_A}{M_w} \approx \frac{1.0 \times (64 \times 10^{-24}) \times 6.022 \times 10^{23}}{18} \approx 2140 \quad (13)$$

*This validates why we insert roughly 2000 molecules.*

## 2. Geometric Construction:

```
1 |     read_data C60_bonded.data add append shift -8.0 0.0 0.0
2 |     read_data C60_bonded.data add append shift 8.0 0.0 0.0
3 |     read_data C60_bonded.data add append shift 0.0 0.0 10.0
```

**Linear Algebra Operation:** We perform three separate read operations. For the  $k$ -th read, every coordinate vector  $\mathbf{x}$  is transformed:

$$\mathbf{x}_{final} = \mathbf{x}_{initial} + \mathbf{T}_k \quad (14)$$

The configuration vectors  $\mathbf{T}$  form a triangle in 3D space:

$$\mathbf{T}_1 = (-8, 0, 0), \quad \mathbf{T}_2 = (8, 0, 0), \quad \mathbf{T}_3 = (0, 0, 10)$$

This specific geometry allows us to study aggregation kinetics (distances between centers  $\approx 16\text{\AA}$ ).

### 3. Lattice Insertion:

```
1 |     lattice sc 3.1
2 |     create_atoms 0 box mol h2omol ...
```

*Math:* Generates a grid  $\{(ix, jy, kz)\}$  with spacing  $\Delta = 3.1\text{\AA}$ . This guarantees initial density  $\rho_{init} \approx (18 \text{ g/mol})/(3.1\text{\AA})^3 \approx 0.99 \text{ g/cm}^3$ .

### 4. Overlap Removal:

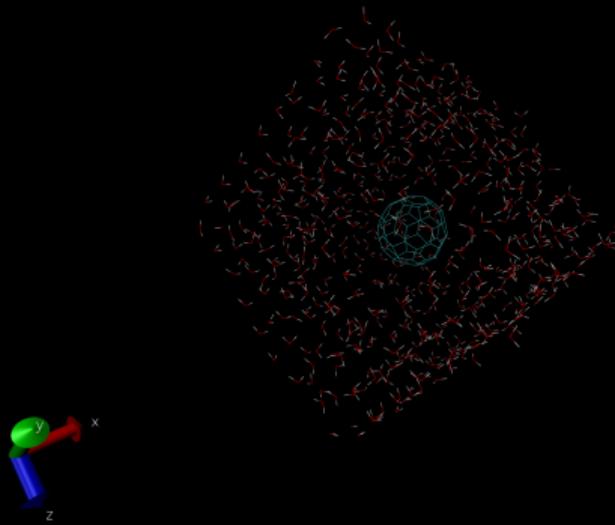
```
1 |     delete_atoms overlap 7.0 water_0 c60_group mol yes
```

*Condition:* Let  $S_C$  be the set of Carbon atoms and  $S_O$  be Oxygen atoms.

$$\text{Delete } O_j \text{ if } \min_{i \in S_C} \|\mathbf{r}_{O_j} - \mathbf{r}_{C_i}\| < 7.0\text{\AA} \quad (15)$$

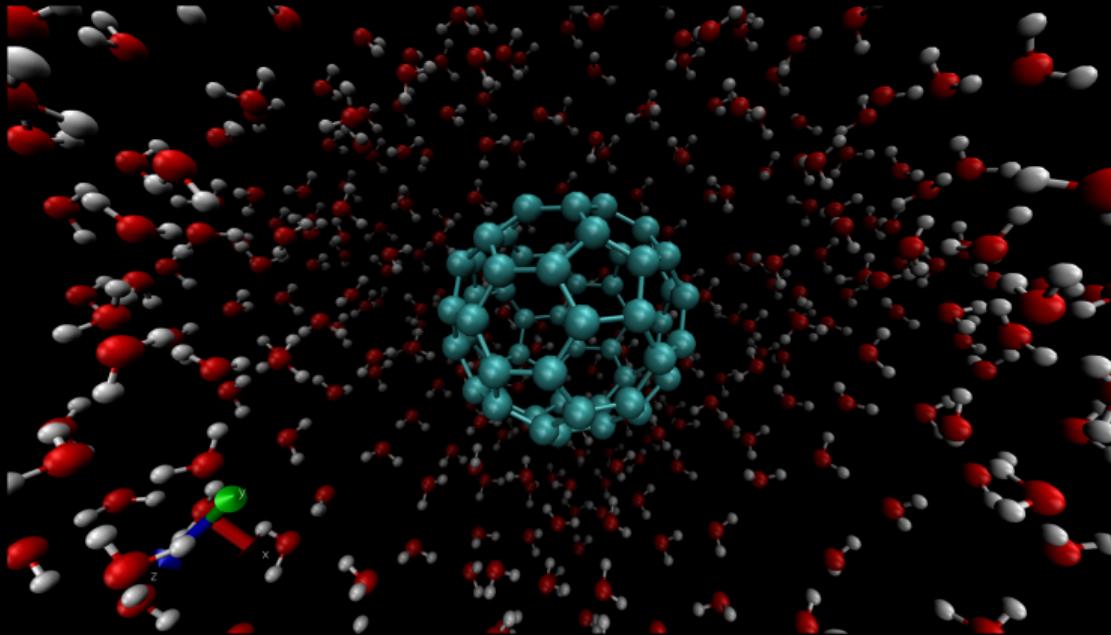
*Why 7.0 \AA?* The C60 radius is  $\sim 3.5\text{\AA}$ . The VdW radius of water is  $\sim 1.6\text{\AA}$ .  $3.5 + 1.6 = 5.1\text{\AA}$ . We add a safety buffer (7.0) to prevent "atomic fusion" (infinite energy) at step 0.

## A Snapshot of the System So Far



**Description:** Three C<sub>60</sub> molecules (gray) surrounded by a dense shell of water molecules (red/white). Overlapping waters have been removed to ensure steric validity.

## A Snapshot of the System So Far



**Description:** Three C<sub>60</sub> molecules (gray) surrounded by a dense shell of water molecules (red/white).

### Generated Artifacts:

1. large\_C60\_solvated.data:
  - ▶ 180 Carbon Atoms (3 Molecules)
  - ▶ ~6000 Water Atoms
  - ▶ Full Topology (Bonds/Angles)
2. \*.xyz: Visual verification.

### Ready for Simulation:

The system is now:

- ▶ Charge Neutral
- ▶ Density optimized ( $\sim 1$  g/cc)
- ▶ Sterically valid (No overlaps)

*Next Step: NPT Equilibration on GPU.*

## Step 4: Building the Large System (1\_build\_large\_C60\_system.lmp)

### Box & Geometry:

```
1 |     region mybox block -20 20 -20 20 -20 20  
2 |     create_box 3 mybox ...
```

**Math:** Define Domain  $\Omega = \{(x, y, z) | -20 \leq x, y, z \leq 20\}$ . Volume  $V = 40^3 = 64,000 \text{ \AA}^3$ .

### Particle Placement:

```
1 |     read_data C60_bonded.data add append shift -8.0 0.0 0.0  
2 |     read_data C60_bonded.data add append shift 8.0 0.0 0.0  
3 |     read_data C60_bonded.data add append shift 0.0 0.0 10.0
```

**Math:** Affine Transformation (Translation) for molecule  $k$ :

$$\mathbf{r}_{i,new}^{(k)} = \mathbf{r}_{i,original} + \mathbf{T}_k \quad \text{where } \mathbf{T}_1 = \begin{bmatrix} -8 \\ 0 \\ 0 \end{bmatrix}, \dots \quad (16)$$

## Step 4: Solvation & Smart Deletion

```
1 |     lattice sc 3.1
2 |     create_atoms 0 box mol h2omol ...
3 |     delete_atoms overlap 7.0 water_0 c60_group mol yes
```

### Solvation Algorithm:

1. Generate solvent grid points  $P_{grid}$  via Simple Cubic (sc) lattice.
2. Insert molecule template at each point.

**Exclusion Logic (Mathematical Condition):** Let  $S_{C60}$  be the set of all Carbon atoms. A water molecule  $W$  (centered at  $\mathbf{r}_O$ ) is deleted if:

$$\exists \mathbf{c} \in S_{C60} : ||\mathbf{r}_O - \mathbf{c}|| < R_{cutoff} \quad (17)$$

Here,  $R_{cutoff} = 7.0 \text{ \AA}$ . This prevents "nuclear fusion" (infinite energy overlaps) at  $t = 0$ .

## Force Field Parameters

```
1 |      # C60-Water Interactions (The Variable)
2 |      pair_coeff 1 2 0.05 3.2 # C-O Interaction
```

**The Study Variable  $\epsilon_{CO}$ :** The total potential energy is:

$$U_{total} = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{i < j} \left[ 4\epsilon_{ij} \left( \frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6} \right) + \frac{Cq_i q_j}{r_{ij}} \right] \quad (18)$$

We vary  $\epsilon_{C-O}$  (1-2 interaction).

- ▶ **Low  $\epsilon$ :** Carbon is hydrophobic (weak attraction).
- ▶ **High  $\epsilon$ :** Carbon is hydrophilic (strong attraction).

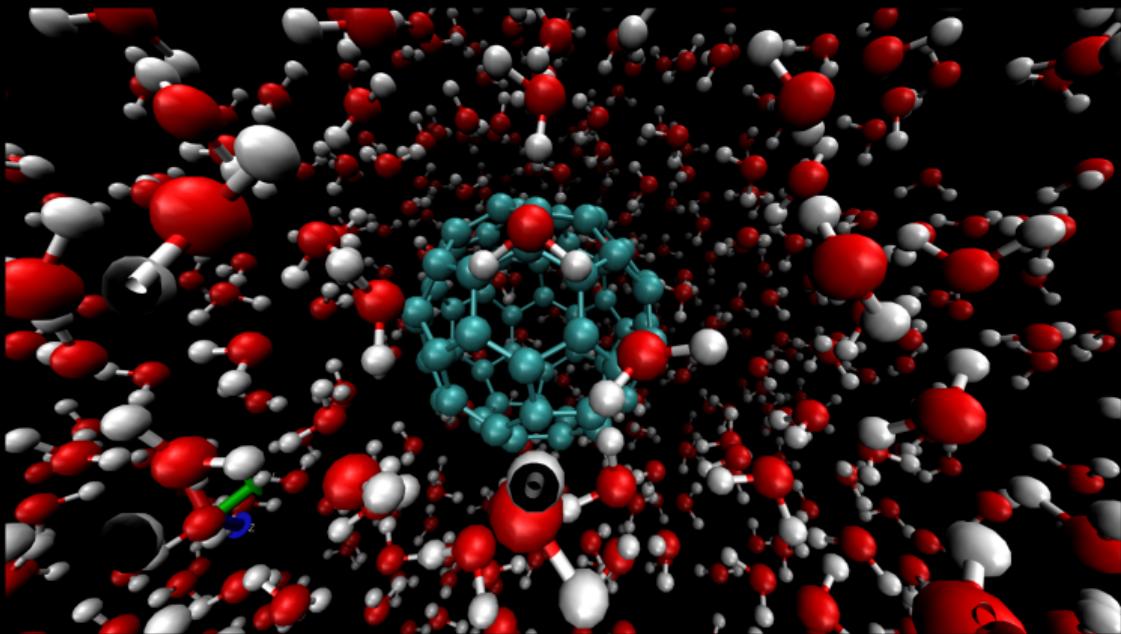
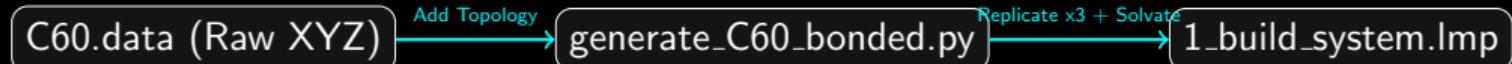


Figure: Final Solvated System: 1 C<sub>60</sub> molecules (gray) surrounded by water (red/white).

## Summary of Setup



Ready for Production

**Outcome:** A fully solvated, charge-neutral, topologically valid system prepared for GPU-accelerated equilibration.

# The Main MD Script

Script: 2\_equilibrium\_version\_2\_w\_minimization.lmp

Solving the "Gas Phase Explosion" via Hybrid Hardware & Numerical Stabilization

**The Hardware Bottleneck:** Standard pppm/gpu is fast but unstable during high-energy overlaps. pppm/tip4p (CPU) is stable but slow.

```
1 |     units real
2 |     atom_style full
3 |     # GPU Acceleration for Force Calculation
4 |     package gpu 1
5 |     pair_style lj/cut/tip4p/long/gpu 2 3 2 1 0.1546 12.0
```

### Strategy:

1. **Short Range Forces (LJ):** Offloaded to GPU ( $O(N)$  neighbors).
2. **Long Range Coulomb:**
  - ▶ During Minimization: CPU (Stability).
  - ▶ During MD: GPU (Speed).

## 1. Initialization & Boundary Conditions

```
1 |     units real
2 |     atom_style full
3 |     boundary p p p
4 |     atom_modify map array
```

### Physical Interpretations:

- ▶ units real:
  - ▶ Energy: kcal/mol
  - ▶ Distance: Ångstroms ( $10^{-10}$  m)
  - ▶ Time: femtoseconds ( $10^{-15}$  s)
- ▶ atom\_style full: Required for **TIP4P water**. We need:

$$\{ID, MolID, Type, q, x, y, z\}$$

Standard ‘atomic’ style lacks Charge ( $q$ ) and Molecule ID.

- ▶ boundary p p p: Periodic Boundary Conditions (PBC).

$$\mathbf{r}_{ij} = \min(|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}L|) \quad \mathbf{n} \in \mathbb{Z}^3 \quad (19)$$

Simulates an infinite bulk fluid using a finite box.

## 2. Parallel Communication Safety

```
1 |     comm_modify cutoff 18.0
```

**Why modify the cutoff?** LAMMPS decomposes the domain into spatial chunks (processors). Atoms near boundaries are "Ghost Atoms".

**The Constraint:**

$$R_{comm} \geq R_{cut} + R_{skin} \quad (20)$$

- ▶  $R_{cut}$  (Coulomb) = 12.0 Å.
- ▶  $R_{skin}$  (Neighbor) = 2.0 Å.
- ▶ Minimum Safety = 14.0 Å.

**Why 18.0 Å?** We anticipate large volume fluctuations during the pressure ramp (Gas → Liquid transition). If the box expands rapidly, atoms move far. 18.0 Å prevents "Lost Atom" errors during the 169,000 atm spike scenarios.

### 3. GPU-Accelerated Force Field

```
1 | pair_style lj/cut/tip4p/long/gpu 2 3 2 1 0.1546 12.0
```

**The TIP4P/2005 Model:** Reference: Abascal, J. L., & Vega, C. (2005). *J. Chem. Phys.*, 123(23).

The 0.1546 parameter is the distance  $d_{OM}$  of the massless charge site  $M$ .

- ▶ /gpu: Offloads the  $O(N)$  Lennard-Jones loop to CUDA cores.
- ▶ /long: Signals that Coulombics are handled by KSpace (PPPM).

#### Interaction Types:

- ▶ Type 1: Carbon (C60)
- ▶ Type 2: Oxygen (Water) - Central LJ site
- ▶ Type 3: Hydrogen (Water) - Charge only

#### 4. Long-Range Solver (The Fix)

```
1 |     kspace_style pppm/gpu 1.0e-4  
2 |     kspace_modify mesh 50 50 50 order 5 force/disp/real 0.8
```

**Particle-Particle Particle-Mesh (PPPM):** A variant of Ewald Summation. Maps charges to a 3D grid to solve Poisson's equation via FFT.

**Parameter Choices:**

- ▶ **Tolerance ( $10^{-4}$ ):** Relative error in forces. Standard for liquid water.
- ▶ **Mesh 50 50 50:** Locked to prevent "grid jitter" artifacts.

$$\text{Spacing} \approx 40\text{\AA}/50 = 0.8\text{\AA} \quad (\text{Rule of thumb: } < 1.0\text{\AA})$$

- ▶ **Order 5:** B-spline interpolation order. Higher order = smoother forces, less energy drift, but higher communication cost. 5 is high-precision (standard is 3 or 4).

## 5. Pair Coefficients (The Science)

```
1 # Water (TIP4P/2005)
2 pair_coeff 2 2 0.1852 3.1589    # 0-0 epsilon(kcal/mol) sigma(Ang)
3
4 # Carbon (C60)
5 pair_coeff 1 1 0.07 3.4          # C-C (Standard Aromatic)
6
7 # Cross Interaction (The Study Variable)
8 variable epsilon_co equal ${EPSILON_CO}
9 pair_coeff 1 2 ${epsilon_co} 3.2
```

**Lorentz-Berthelot Mixing Rules (Implicit):** Usually  $\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}$ . Here,  $\sigma_{CO} = \frac{3.4 + 3.1589}{2} \approx 3.28$ . We use 3.2 explicitly based on literature tuning for carbon nanotubes in water.

**The Variable  $\epsilon_{CO}$ :** Controls the wetting angle.

- ▶  $\epsilon < 0.1$ : Hydrophobic (dewetting).
- ▶  $\epsilon > 0.2$ : Hydrophilic (wetting).

## 6. Neighbor Lists & Binning

```
1 |     neighbor 2.0 bin  
2 |     neigh_modify delay 5 every 1 check yes
```

**Verlet Lists:** To avoid  $O(N^2)$  distance checks, we keep a list of atoms within  $R_{cut} + R_{skin}$ .

$$R_{list} = 12.0 + 2.0 = 14.0 \text{ \AA} \quad (21)$$

**Why bin?** LAMMPS sorts atoms into spatial bins (cells) of size  $\approx R_{list}/2$ .

- ▶ List build becomes  $O(N)$  instead of  $O(N^2)$ .
- ▶ Crucial for our system ( $N \approx 6000$ ) to run efficiently.

**Update Frequency:** ‘check yes’: Only rebuild list if an atom moved  $> R_{skin}/2$  (1.0 Å). ‘delay 5’: Don’t even check for the first 5 steps after a rebuild (optimization).

## 7. Timestep & Integration

```
1 |    timestep 2.0
```

**Integration Stability:** The Velocity-Verlet algorithm error scales as  $O(\delta t^2)$ .

**Why 2.0 fs? (Standard is 1.0 fs)**

- ▶ **Restriction:** The fastest motion in water is the O-H bond vibration ( $\tau \approx 10$  fs). To resolve it, we need  $\delta t \approx 1$  fs.
- ▶ **Solution:** We use fix shake to freeze O-H bonds.
- ▶ **Result:** The fastest remaining motion is bond angle bending / LJ collisions. This allows extending  $\delta t \rightarrow 2.0$  fs safely.
- ▶ **Benefit:** The simulation runs **2x faster**.

## Critical Fix #1: The PPPM Grid Lock

**The Root Cause of Failure:** Previously, the simulation exploded ( $P > 169,000 \text{ atm}$ ).

*Reason:* The Particle-Particle Particle-Mesh (PPPM) solver auto-resized the FFT grid when the box volume fluctuated, causing a discontinuous jump in Potential Energy.

```
1 | kspace_style pppm/gpu 1.0e-4
2 | # FIX: Lock the grid dimensions
3 | kspace_modify mesh 50 50 50 order 5
```

**Mathematics of PPPM:** The electrostatic energy is split (Ewald Summation):

$$E = E_{real} + E_{reciprocal} + E_{self} + E_{surface} \quad (22)$$

$E_{reciprocal}$  is solved via Fast Fourier Transform (FFT) on a mesh.

$$\text{Grid Spacing } \Delta x = \frac{L_x}{N_{grid}}$$

If  $N_{grid}$  changes discretely while  $L_x$  is changing,  $\Delta E \rightarrow \infty$ .

**Solution:** Force  $N_{grid} = (50, 50, 50)$  constant.

## Stage 0: Gentle Minimization (CPU)

```
1 |     # Switch to CPU pair style for stability
2 |     pair_style lj/cut/tip4p/long ...
3 |     minimize 0.0 10.0 10 100
```

**Mathematical Objective:** Find the local minimum of the Potential Energy Surface (PES).

$$\nabla U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathbf{0} \quad (23)$$

We use the Conjugate Gradient method. We are NOT solving equations of motion (time is frozen). We simply move atoms along the force vector:

$$\mathbf{r}_{new} = \mathbf{r}_{old} + \gamma \mathbf{F} \quad (24)$$

*Purpose:* Resolve "bad contacts" (overlaps) created during the build phase ( $r < \sigma$ ) which would cause infinite forces  $\mathbf{F} \rightarrow \infty$  in MD.

## Stage 1: NVT Thermalization

```
1 |      timestep 2.0
2 |      velocity all create 300.0 ... dist gaussian
3 |      fix mynvt all nvt temp 300.0 300.0 100.0
```

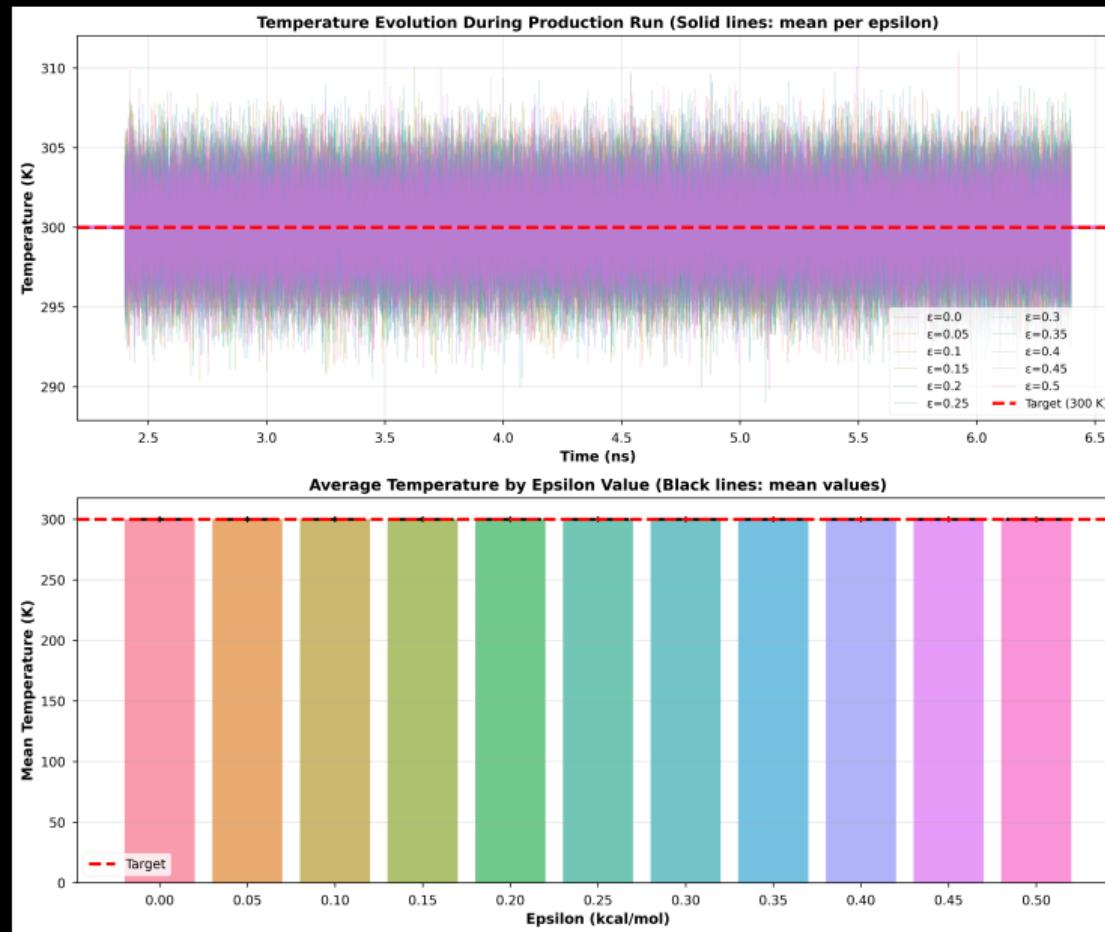
1. **Timestep ( $dt = 2.0 \text{ fs}$ )**: Possible because SHAKE removes high-frequency O-H bond vibrations ( $f \approx 10^{14} \text{ Hz}$ ). Speedup:  $2\times$ .

2. **Velocity Initialization**: Assigns velocities from Maxwell-Boltzmann distribution:

$$P(v) = 4\pi \left( \frac{m}{2\pi k_B T} \right)^{3/2} v^2 \exp\left(\frac{-mv^2}{2k_B T}\right) \quad (25)$$

3. **NVT Ensemble (Canonical)**: Constant  $N$  (atoms),  $V$  (volume),  $T$  (temp). The barostat is OFF. Pressure will be high ( $\sim 800 \text{ atm}$ ) because we packed atoms into a fixed box.

## (Let's see the Temperature Equilibration)



## Stage 1.5: Pre-Equilibration (The Pressure Bridge)

**The Logic:** In the failed run, NVT (Fixed V) → NPT (1 atm) caused a shock. *Physics:* Water has a high Bulk Modulus ( $K \approx 2.2$  GPa). A tiny density mismatch results in massive pressure ( $P \approx 10^5$  atm).

```
1 |     # Fix #2: Target 100 atm, High Damping
2 |     fix mynpt_pre all npt temp 300 300 100 iso 100 100 5000
```

### Parameter breakdown:

- ▶ iso 100 100: We target an intermediate pressure. This allows the box to expand, but effectively "pushes back" to prevent the water from flashing into gas.
- ▶ Pdamp 5000: The barostat relaxation time  $\tau_P$ .

$$\dot{V} \propto \frac{1}{\tau_P} (P_{inst} - P_{target}) \quad (26)$$

A large  $\tau_P$  (5000 fs) makes the box volume react *slowly*, preventing explosive expansion.

## Stage 2: The Pressure Ramp (Fix #3)

### The Command:

```
1 | fix mynpt_ramp all npt ... iso 100.0 1.0 2000.0 drag 2.0
```

**What is happening?** We linearly interpolate the target pressure  $P_{ext}(t)$  from 100 atm to 1 atm over 50,000 steps.

**Mathematics of ‘drag 2.0’:** The volume equation of motion in the Nosé-Hoover-Andersen barostat includes a fictitious “piston mass”  $W$ . The ‘drag’ adds a friction term  $\Gamma$ :

$$W\ddot{V} = V(P_{int} - P_{ext}) - \underbrace{\Gamma\dot{V}}_{\text{Drag Force}} \quad (27)$$

*Why?* Without drag, the box volume  $V(t)$  oscillates like a harmonic oscillator (ringing) after a pressure drop. The drag term acts as a critical damper ( $\zeta \approx 1$ ), stabilizing the density smoothly.

## Stage 3: Two-Phase NPT Equilibration

We do not simply "turn on" NPT. We use a refined 2-phase approach.

### Phase 1: Stabilization (400 ps)

```
1 | fix mynpt_gentle all npt/gpu ... iso 1.0 1.0 5000.0
```

*Action:* Uses npt/gpu (offloads virial calc to GPU). Keeps  $\tau_P = 5000$  fs. The system gently settles into the 1 atm basin.

### Phase 2: Convergence (600 ps)

```
1 | fix mynpt_standard all npt/gpu ... iso 1.0 1.0 2000.0
```

*Action:* We tighten  $\tau_P$  to 2000 fs. This allows standard thermal fluctuations of the volume, necessary for correct density sampling.

## Safety Mechanism: Density Checkpoints

Simulations can fail silently. We added logic gates to the script.

```
1 variable current_dens equal density
2 if "${current_dens}<0.85 || ${current_dens}>1.15" then &
3   "print 'ERROR: Density outside expected range'" &
4   "quit 1"
```

**The Gas Phase Detection:** In your previous failed run,  $\rho \rightarrow 0.025 \text{ g/cm}^3$ . This check catches that failure *before* wasting 4 nanoseconds of GPU time on a gas-phase simulation.

## Stage 4: Production - The MTK Correction

```
1 | fix mynpt all npt/gpu temp 300 300 100 iso 1 1 1000 mtk yes
```

**Detailed Math: Martyna-Tobias-Klein (MTK):** Standard Nosé-Hoover generates the NPT distribution only in the thermodynamic limit ( $N \rightarrow \infty$ ). For finite systems, it misses a metric factor in phase space.

MTK adds a correction to the Hamiltonian equations:

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \frac{\mathbf{p}_i \pi_\epsilon}{Q} - \underbrace{\frac{d-1}{dNk_B T} \epsilon \mathbf{p}_i}_{\text{MTK Term}} \quad (28)$$

*Significance:* Ensures that the fluctuations in Volume  $\langle (\delta V)^2 \rangle$  (related to compressibility) are statistically rigorous.

## Analysis: C60-C60 RDF Issue (Sadly I identified it a few hours ago)

**The Problem:** C60 molecules have IDs 1, 2, 3. But they contain 180 atoms total. Standard RDF counts distances between *all* pairs of type 1 (Carbon).

Intra-molecular C-C bonds: 1.4Å (Noise)

Inter-molecular C-C gap: 10 – 20Å (Signal)

**The Fix:**

```
1 |     neigh_modify exclude molecule/intra carbon
2 |     compute rdf_CC all rdf 150 1 1
3 |     neigh_modify exclude none # Restore immediately
```

**Set Theory:** Let  $D = \{||\mathbf{r}_i - \mathbf{r}_j||\}$ . We compute  $D_{inter} = D \setminus \{\text{pairs where } mol(i) == mol(j)\}$ .

## 1. Mean Squared Displacement (MSD):

```
1 |     compute msd_water oxygen msd
```

Used to calculate Self-Diffusion Coefficient ( $D_{water}$ ):

$$D_{water} = \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle \quad (29)$$

*Goal:* See if hydrophobic C60 speeds up water diffusion vs hydrophilic.

## 2. High-Res Thermodynamics:

```
1 |     fix thermo_detailed all ave/time 10 10 100 ...
```

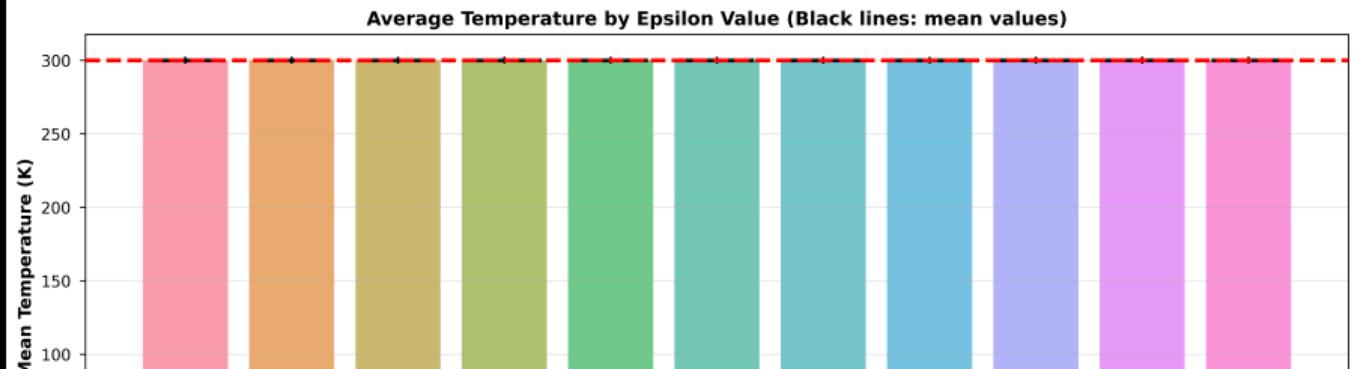
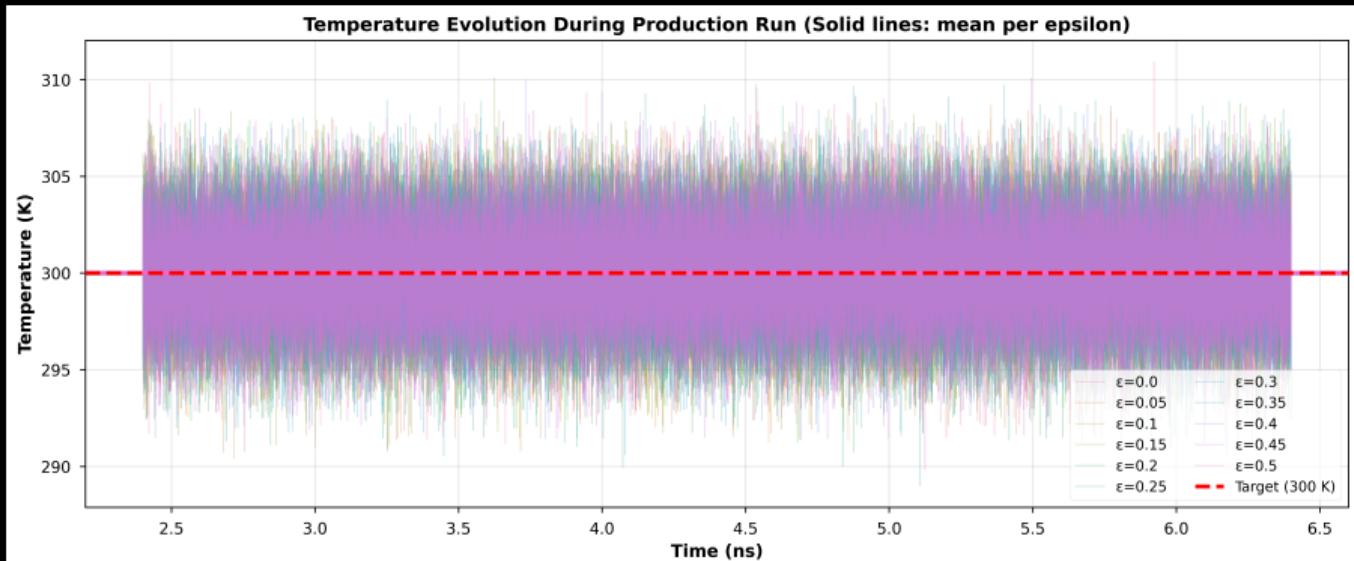
Samples every 0.2 ps. Used to compute heat capacity ( $C_p$ ) and compressibility ( $\kappa_T$ ) via fluctuation formulas:

$$\langle \delta H^2 \rangle = k_B T^2 C_p \quad (30)$$

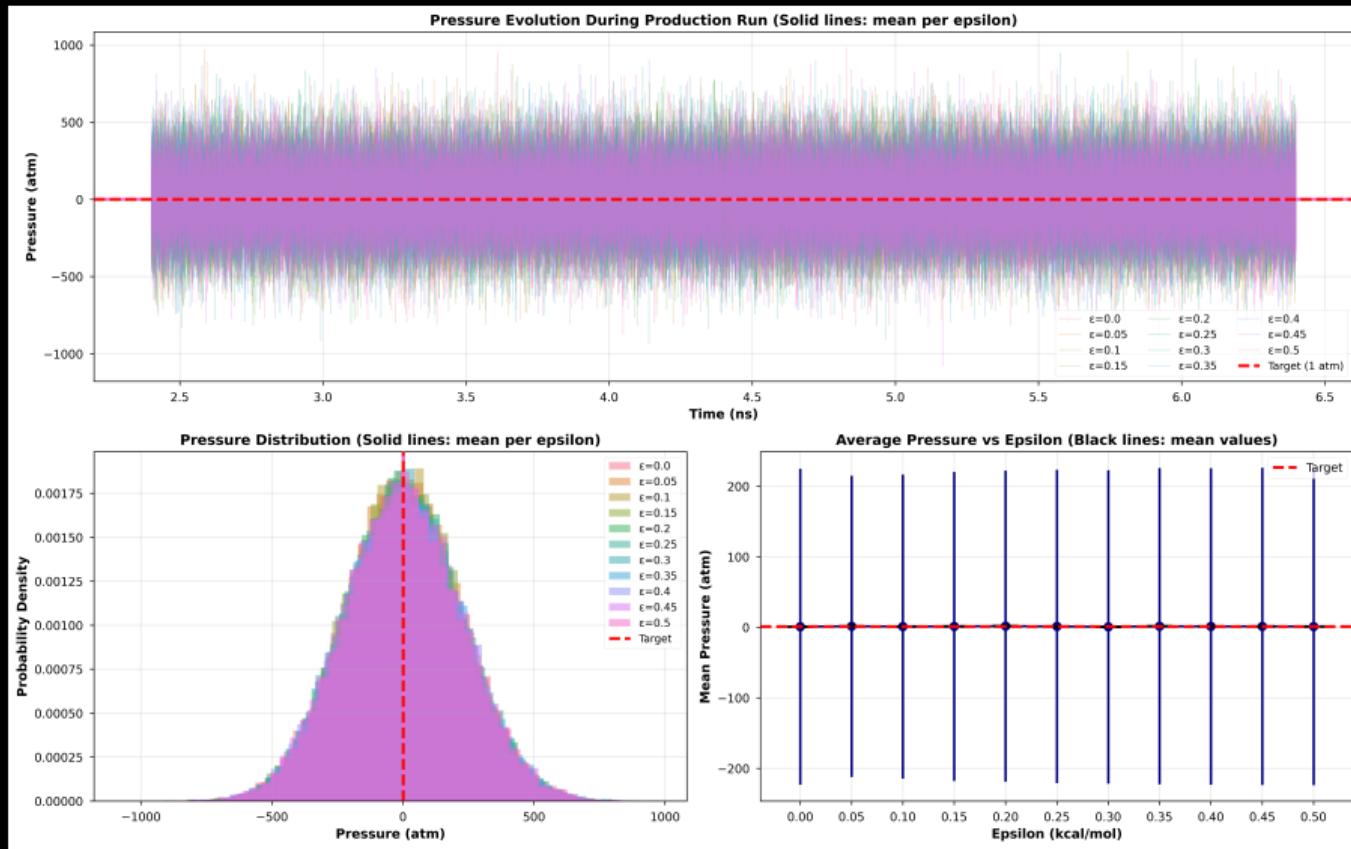
### Ensemble averages & fluctuations

- ▶ Ensemble average:  $\langle X \rangle = \frac{1}{N} \sum_{i=1}^N X_i$
- ▶ Fluctuations:  $\sigma_X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}$
- ▶ Equilibration detection: linear drift  $X(t) = \alpha t + \beta$ , equilibrium when  $\alpha \approx 0$ .
- ▶ Autocorrelation:  $C_X(\tau) = \frac{\langle (X(t) - \langle X \rangle)(X(t+\tau) - \langle X \rangle) \rangle}{\sigma_X^2}$

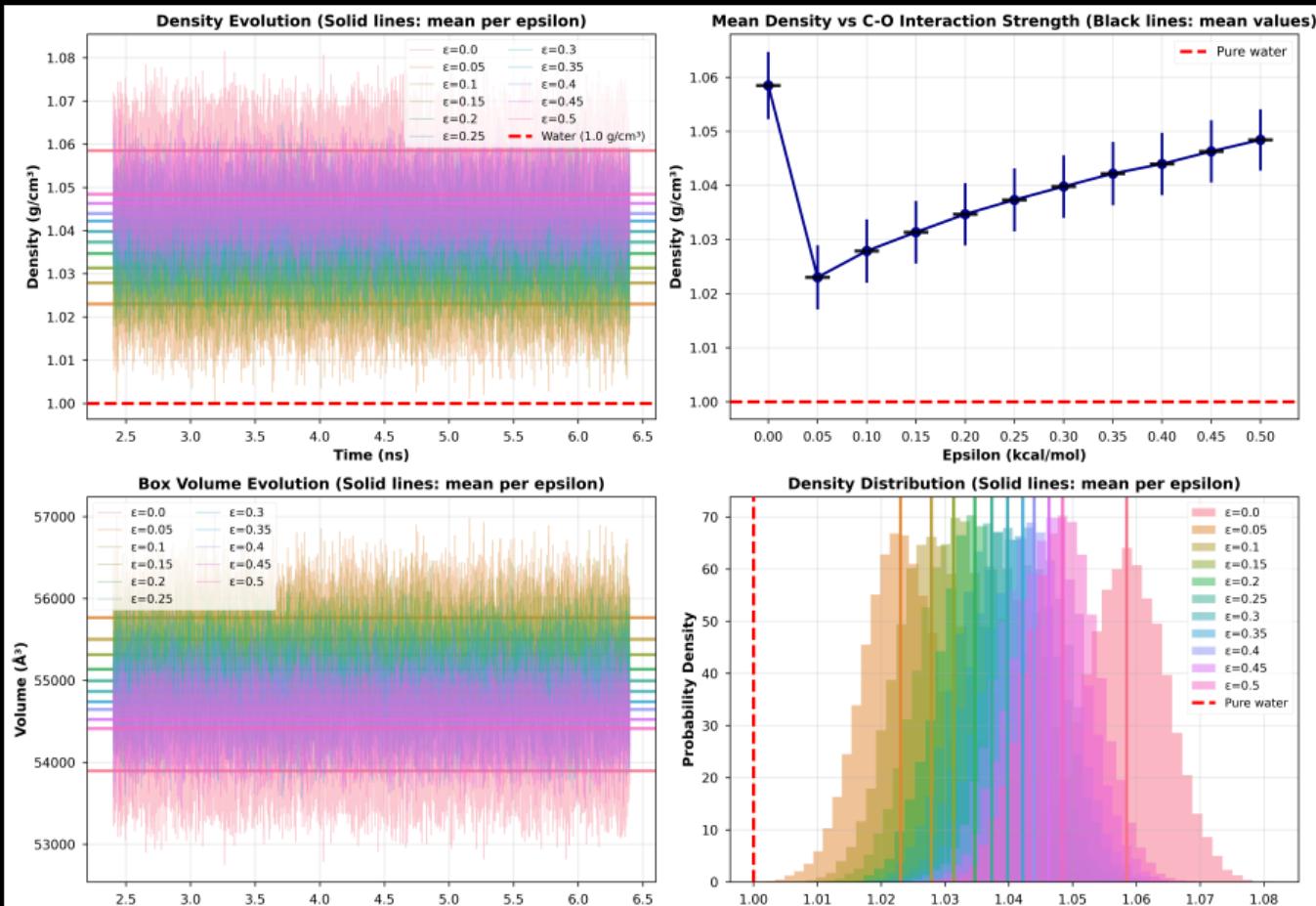
## 01 Temperature Evolution



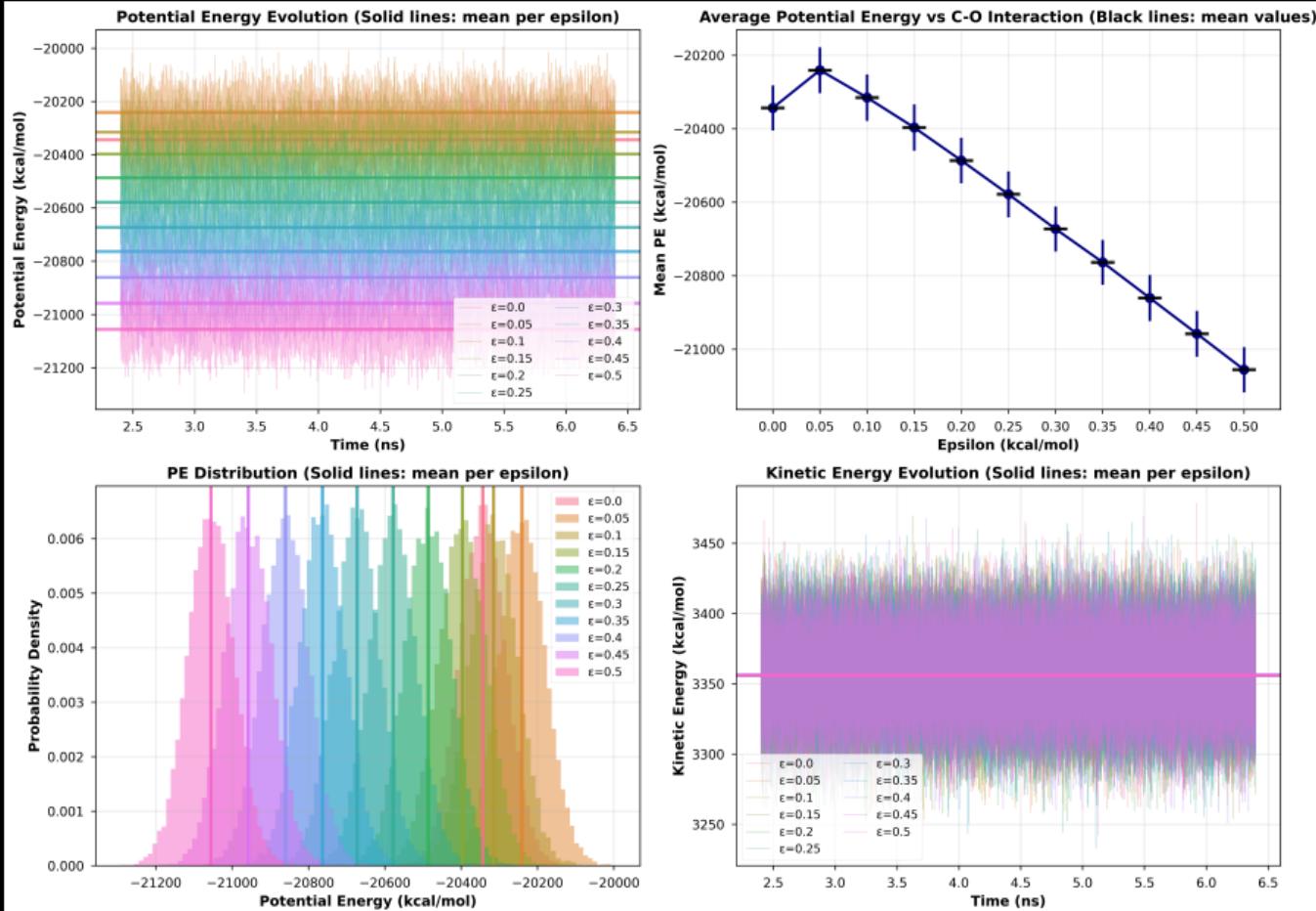
## 02 Pressure Analysis



## 03 Density Analysis

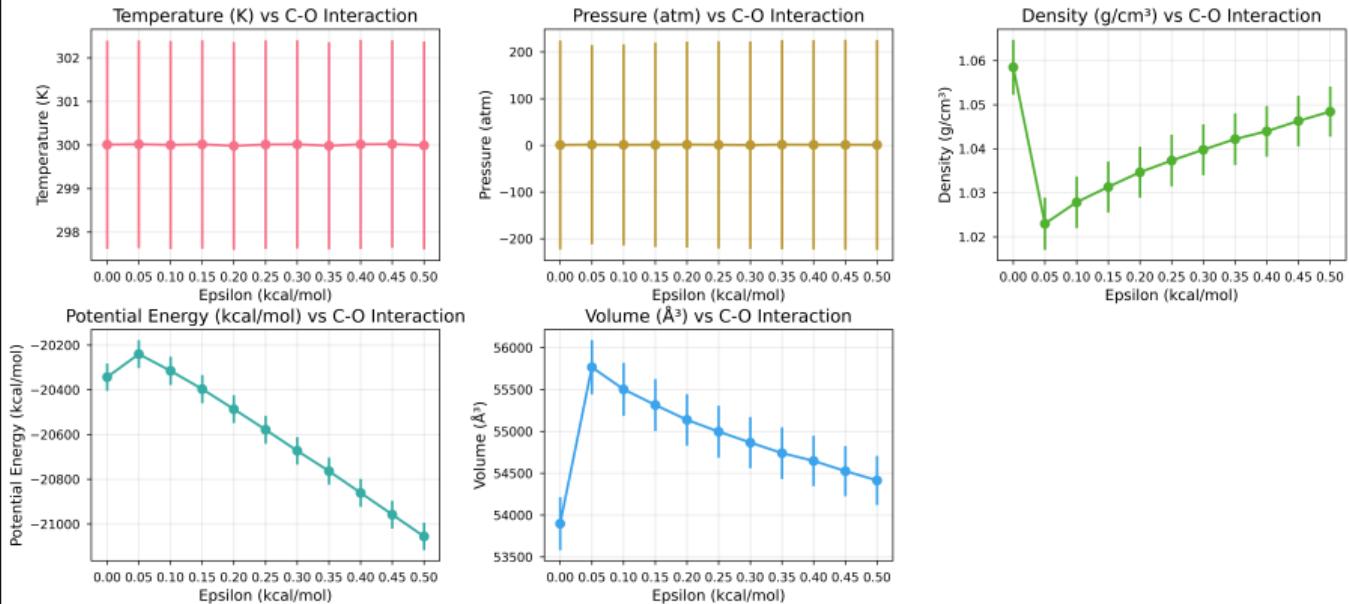


## 04 Energy Analysis



## 05 Comparison Matrix

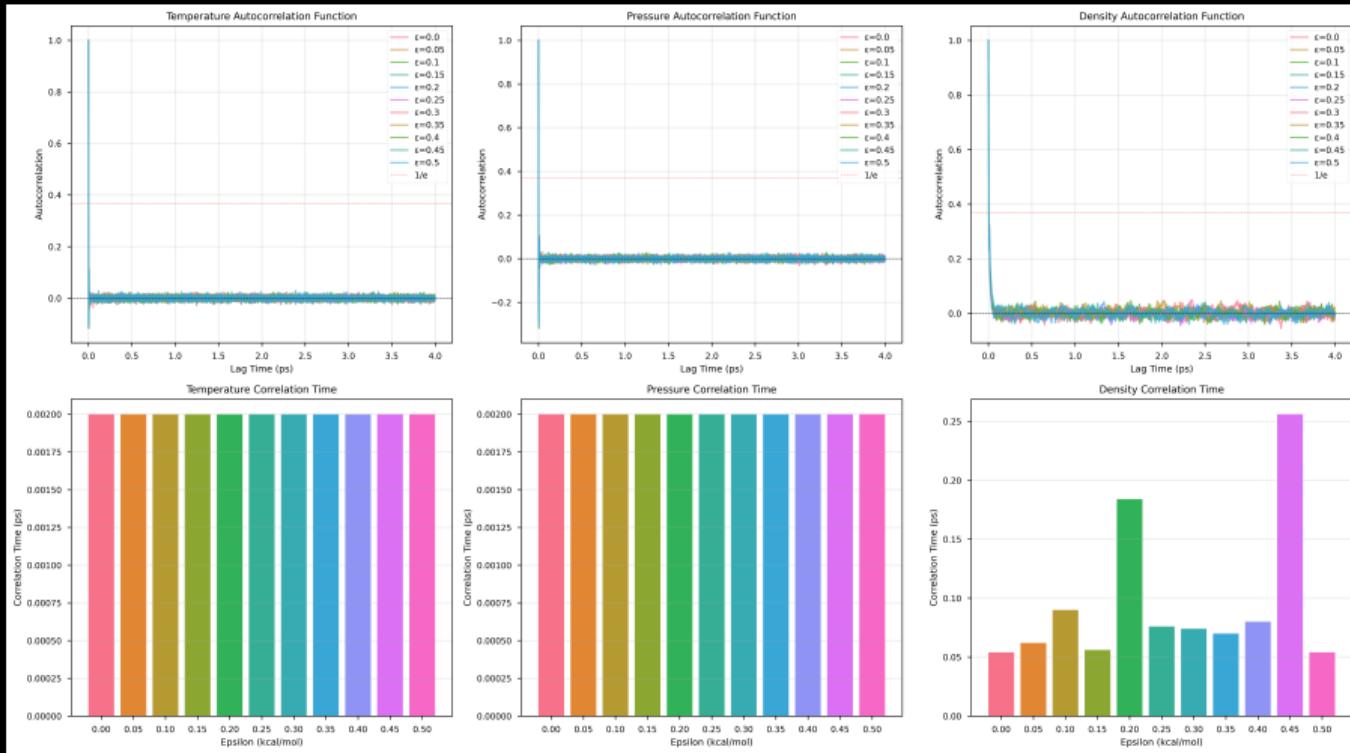
Comprehensive Thermodynamic Comparison Across Epsilon Values



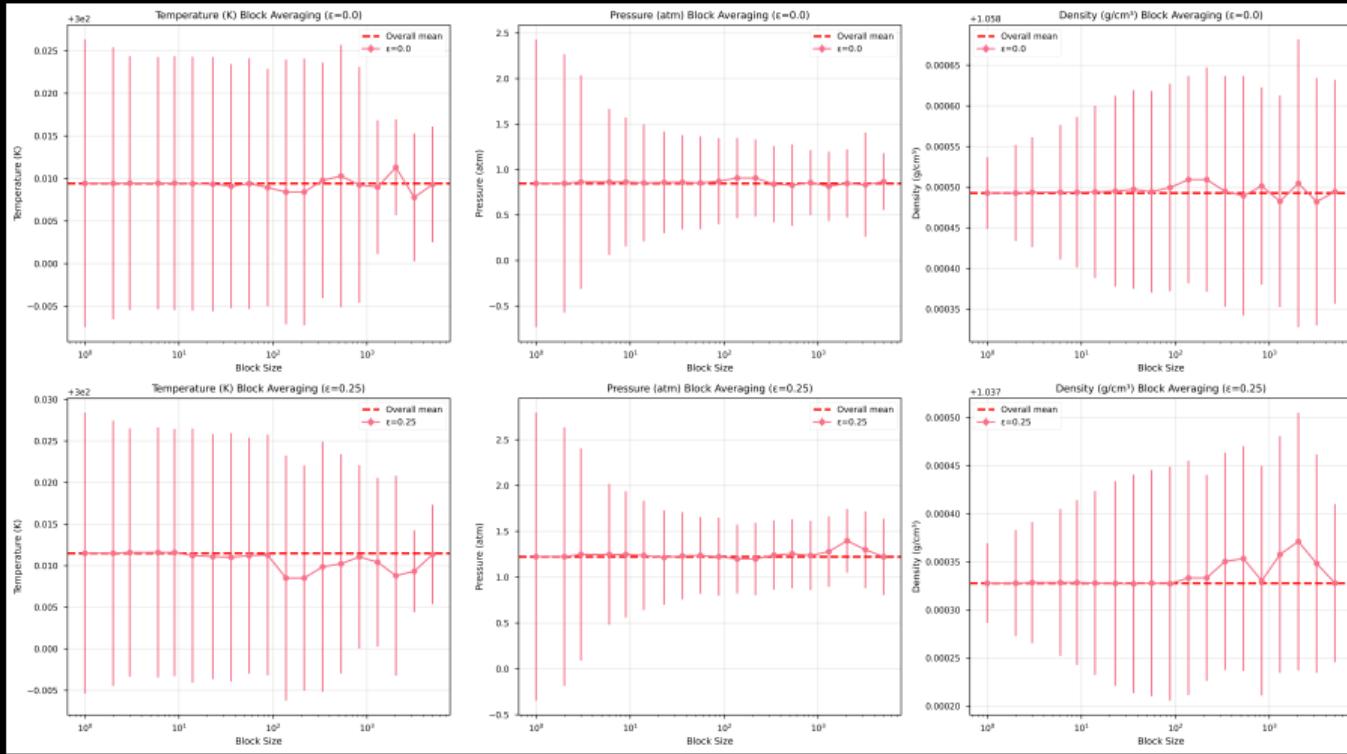
$\epsilon$	Temp (K)	Press (atm)	Dens (g/cm³)	PE (kcal/mol)
0.00	$300.01 \pm 2.39$	$0.8 \pm 223.9$	$1.0585 \pm 0.0062$	$-20343.5 \pm 61.4$
0.05	$300.02 \pm 2.39$	$1.4 \pm 213.5$	$1.0230 \pm 0.0059$	$-20240.7 \pm 62.5$
0.10	$300.00 \pm 2.39$	$1.1 \pm 215.6$	$1.0279 \pm 0.0058$	$-20315.3 \pm 63.2$
0.15	$300.02 \pm 2.39$	$1.3 \pm 219.0$	$1.0314 \pm 0.0058$	$-20396.7 \pm 63.1$
0.20	$299.98 \pm 2.39$	$1.7 \pm 220.5$	$1.0347 \pm 0.0058$	$-20486.4 \pm 62.0$
0.25	$300.01 \pm 2.39$	$1.2 \pm 222.0$	$1.0373 \pm 0.0058$	$-20578.3 \pm 62.3$
0.30	$300.02 \pm 2.39$	$0.6 \pm 222.1$	$1.0398 \pm 0.0058$	$-20672.7 \pm 61.4$
0.35	$299.98 \pm 2.38$	$1.5 \pm 224.1$	$1.0422 \pm 0.0059$	$-20763.3 \pm 60.9$
0.40	$300.02 \pm 2.40$	$1.2 \pm 224.3$	$1.0440 \pm 0.0058$	$-20860.5 \pm 62.6$
0.45	$300.02 \pm 2.38$	$1.3 \pm 224.9$	$1.0463 \pm 0.0057$	$-20957.7 \pm 62.4$

- ▶ Block averaging: divide trajectory into  $M$  blocks, check variance scaling  $\sigma_B^2 \propto 1/B$ .
- ▶ Autocorrelation time and effective sample size via FFT (Wiener-Khinchin).

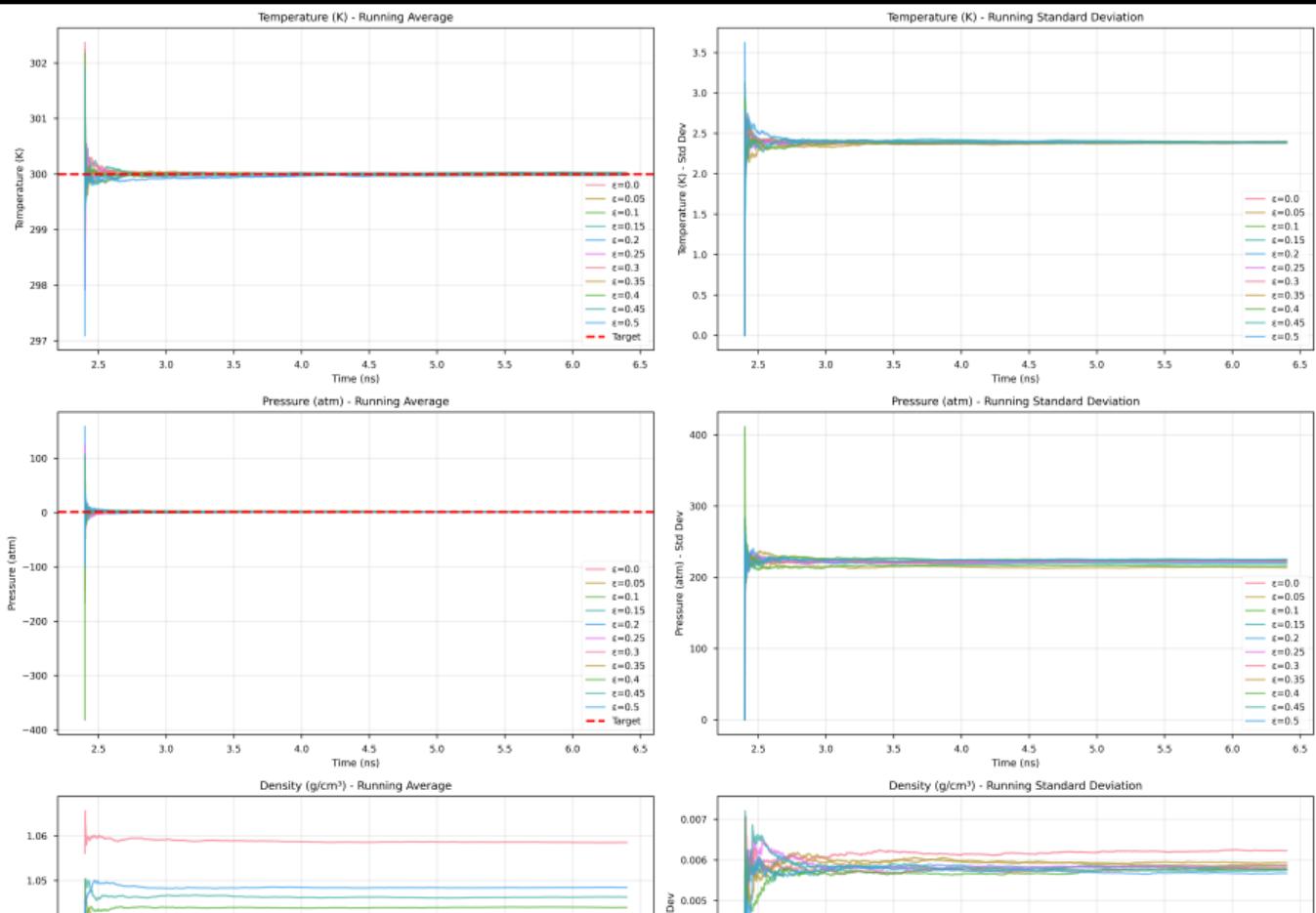
## 06 Autocorrelation Analysis



## 07 Block Averaging

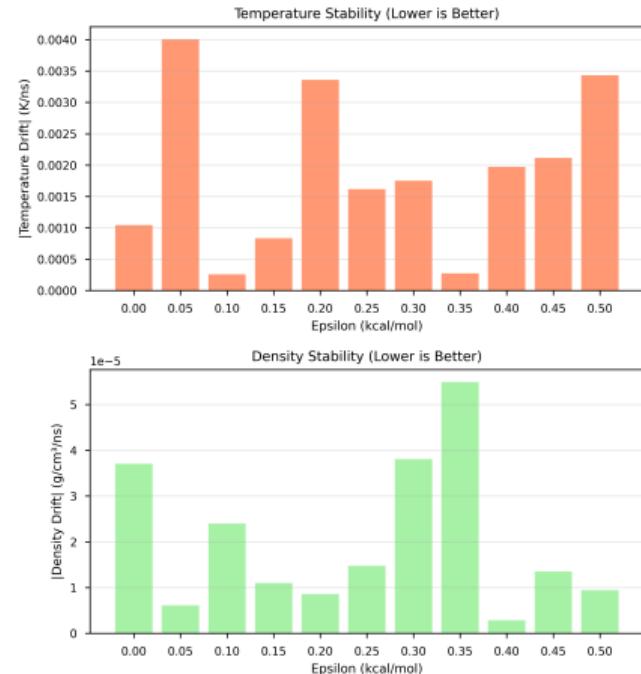
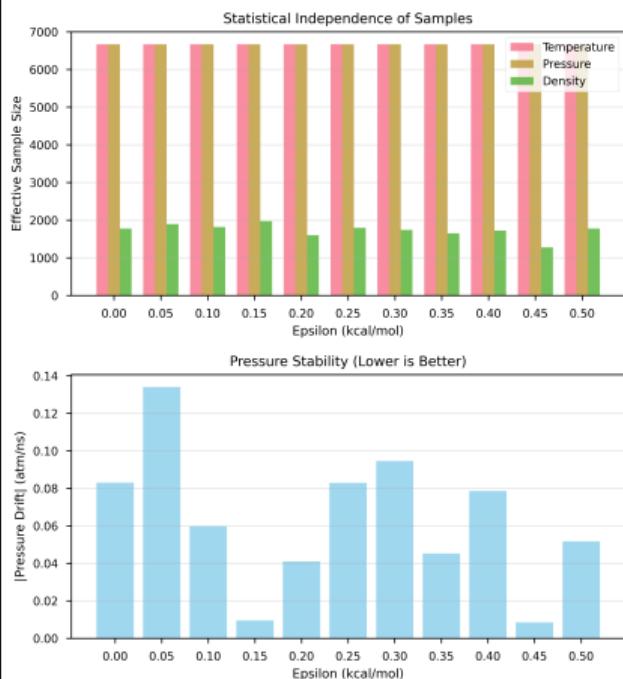


## 08 Running Averages



## 09 Stability Metrics

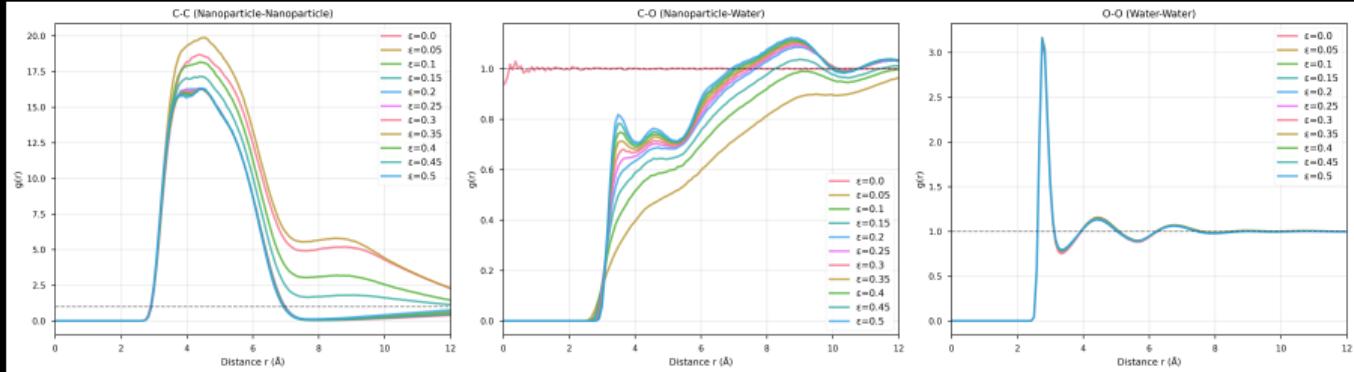
### Equilibration and Stability Analysis



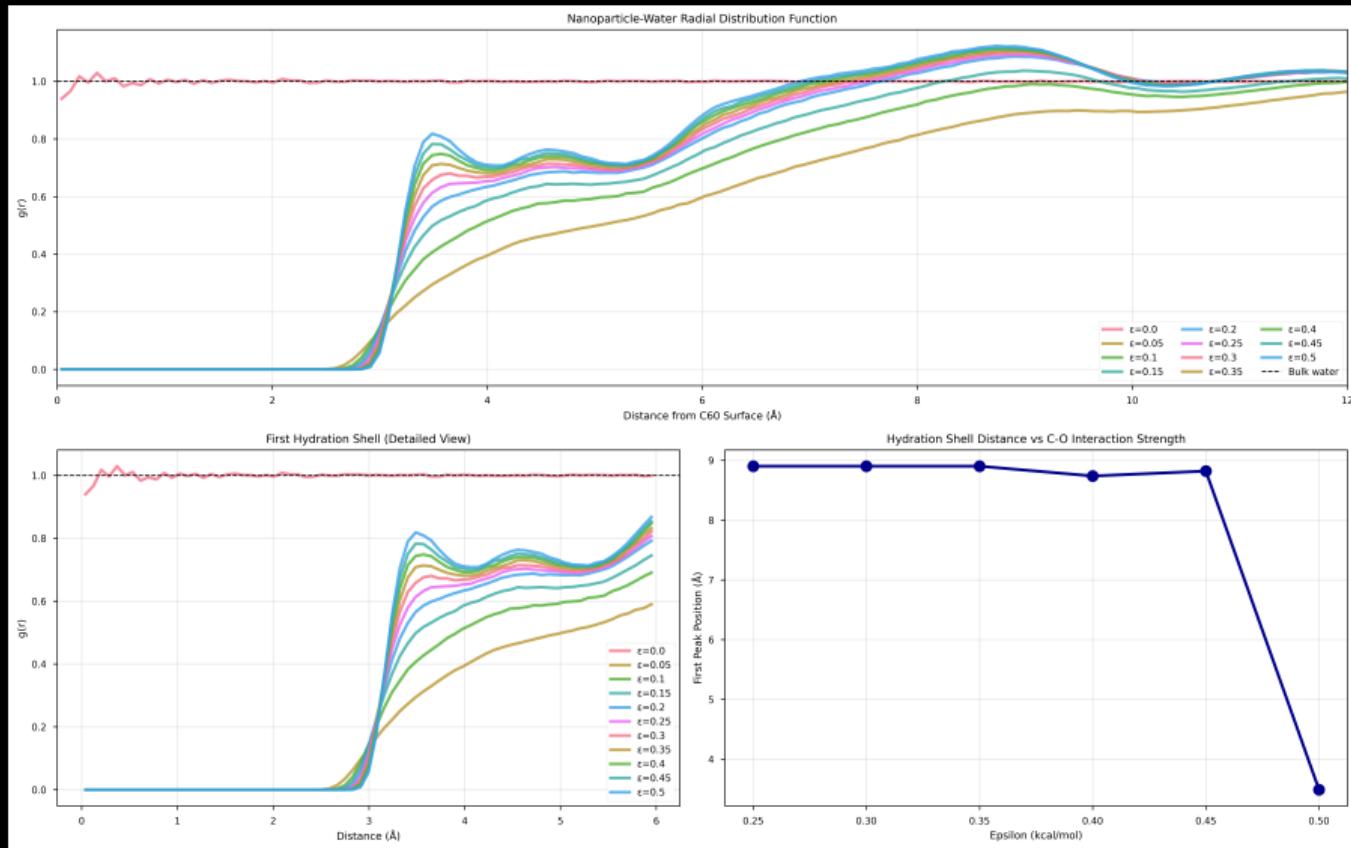
$\epsilon$	$\tau_{\cdot T}$ (ps)	$N_{eff\_T}$	$ \text{Drift}_T $	$\tau_{\cdot P}$ (ps)	$ \text{Drift}_P $	$\tau_{\cdot p}$ (ps)	$ \text{Drift}_p $
0.00	0.0	6667	1.04e-03	0.0	8.30e-02	0.1	3.71e-05
0.05	0.0	6667	4.01e-03	0.0	1.34e-01	0.1	6.11e-06
0.10	0.0	6667	2.58e-04	0.0	5.98e-02	0.1	2.40e-05
0.15	0.0	6667	8.33e-04	0.0	9.47e-03	0.1	1.10e-05

- ▶ Radial distribution:  $g(r) = \frac{1}{4\pi r^2 \rho N} \sum_i \sum_{j \neq i} \delta(r - r_{ij}).$
- ▶ Coordination number:  $N(r_{cut}) = 4\pi \rho \int_0^{r_{cut}} r^2 g(r) dr.$

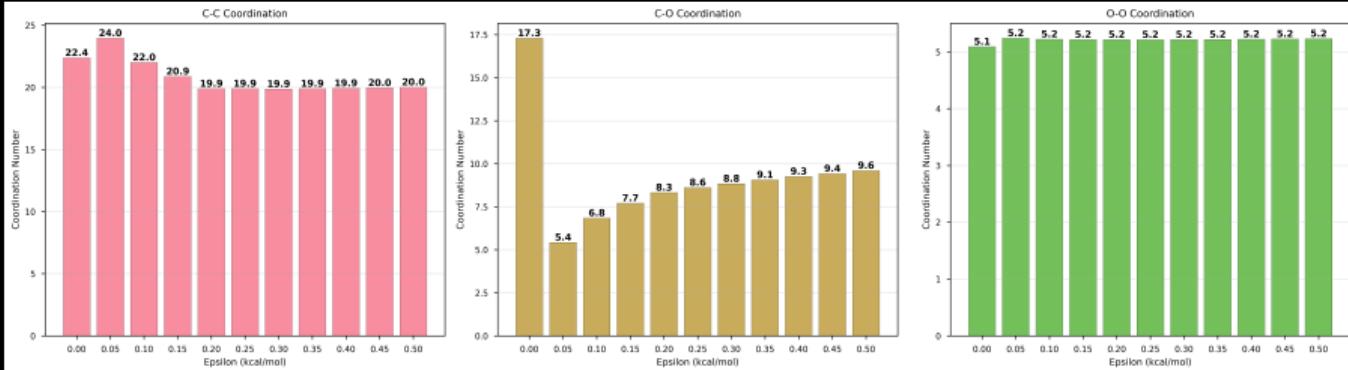
## 10 RDF Comparison



# 11 C-O RDF Detailed

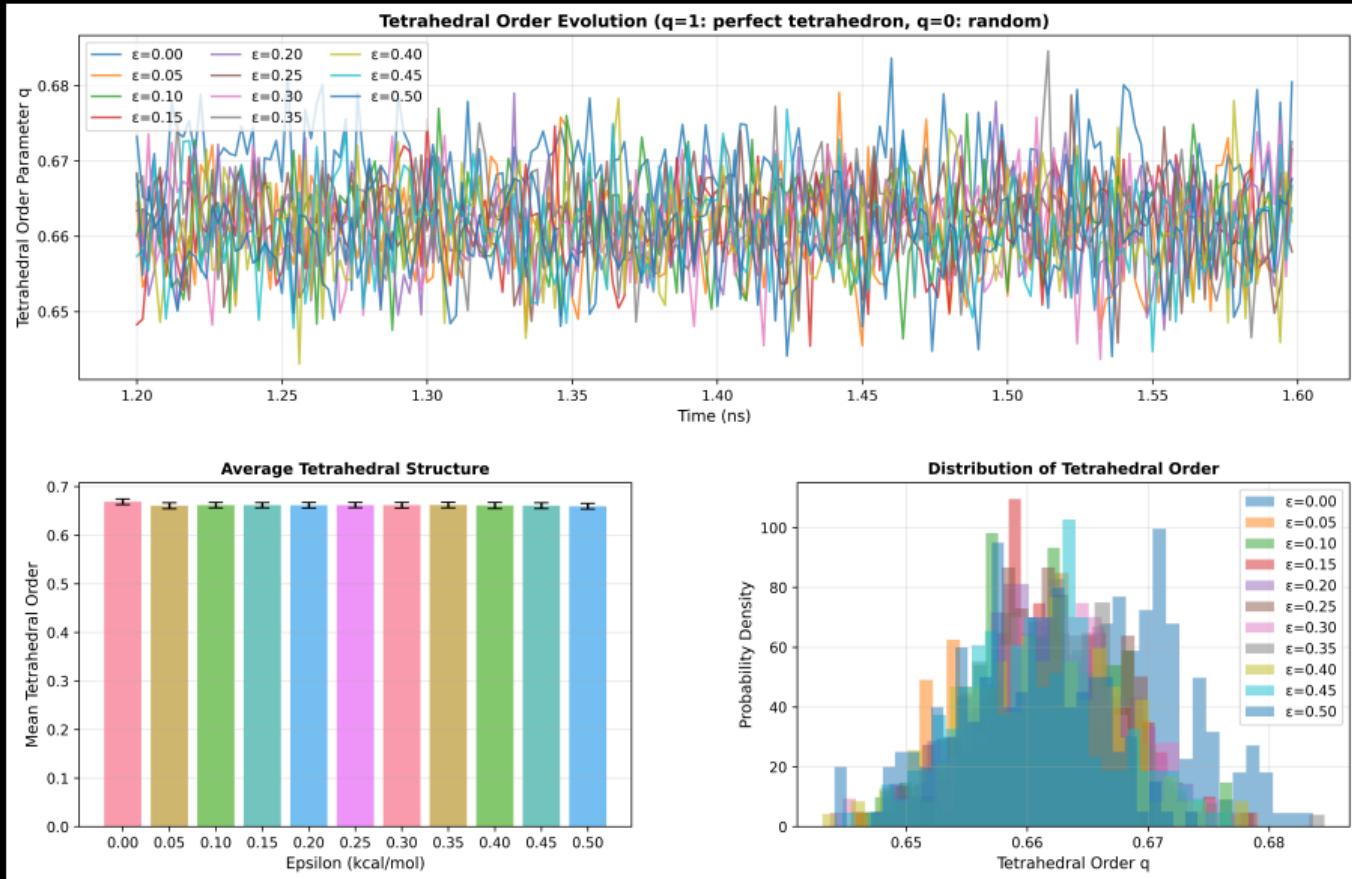


## 12 Coordination Numbers



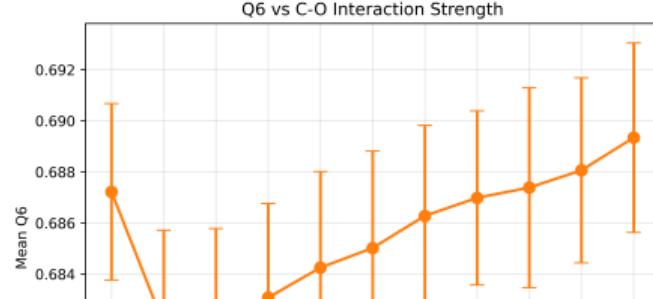
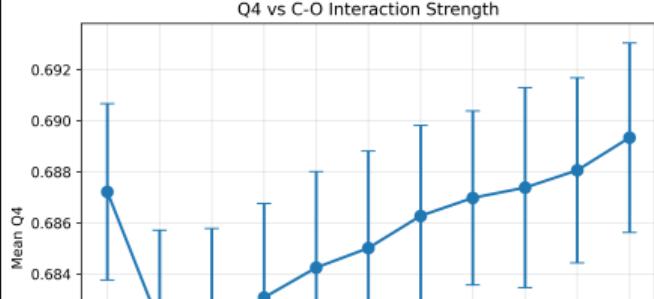
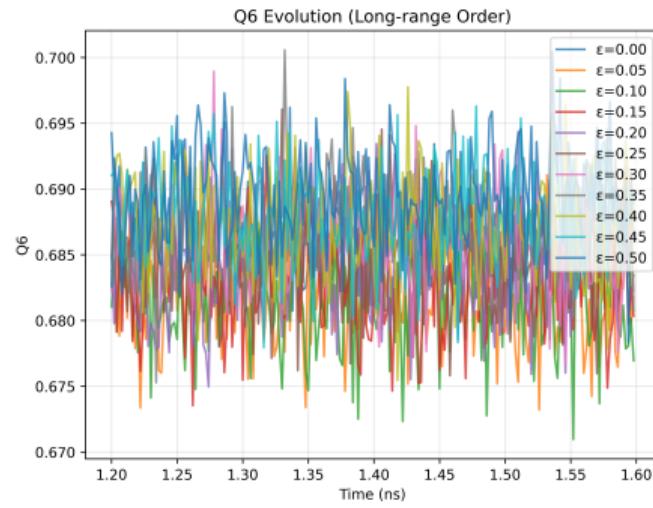
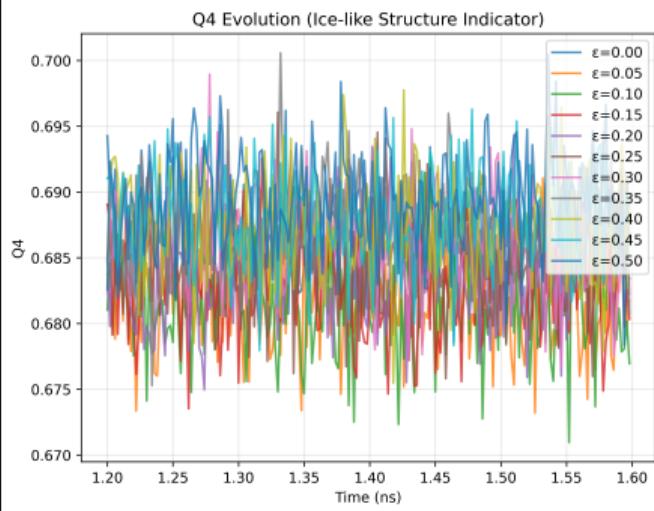
- ▶ Tetrahedral order:  $q = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 (\cos \psi_{jk} + \frac{1}{3})^2$ .
- ▶ Hydrogen bond criteria:  $r_{OO} < 3.5$   
AA,  $\angle H - O \dots O < 30^\circ$ .

## 13 Tetrahedral Order



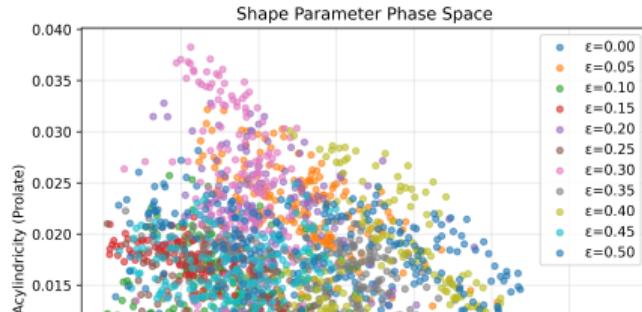
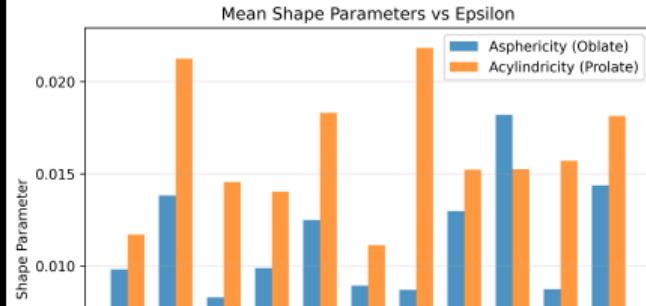
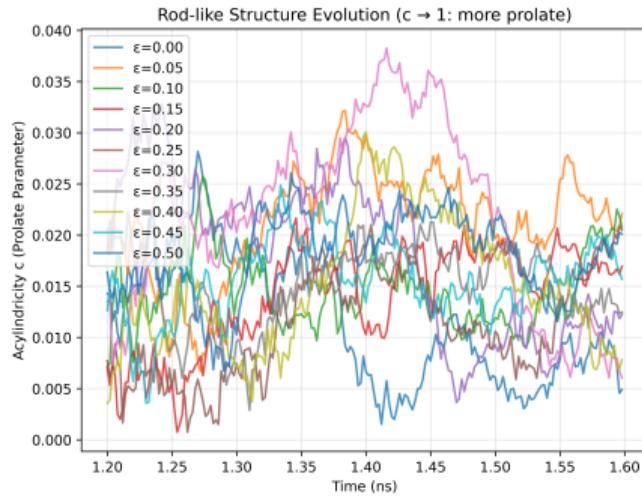
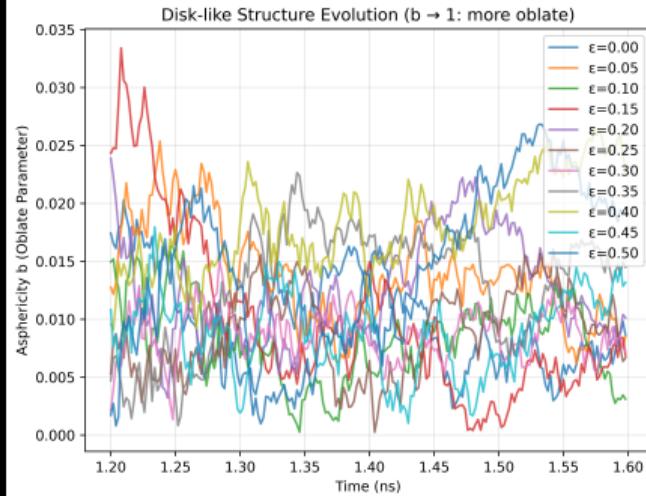
## 14 Steinhardt Order Parameters

Steinhardt Bond-Orientational Order Parameters



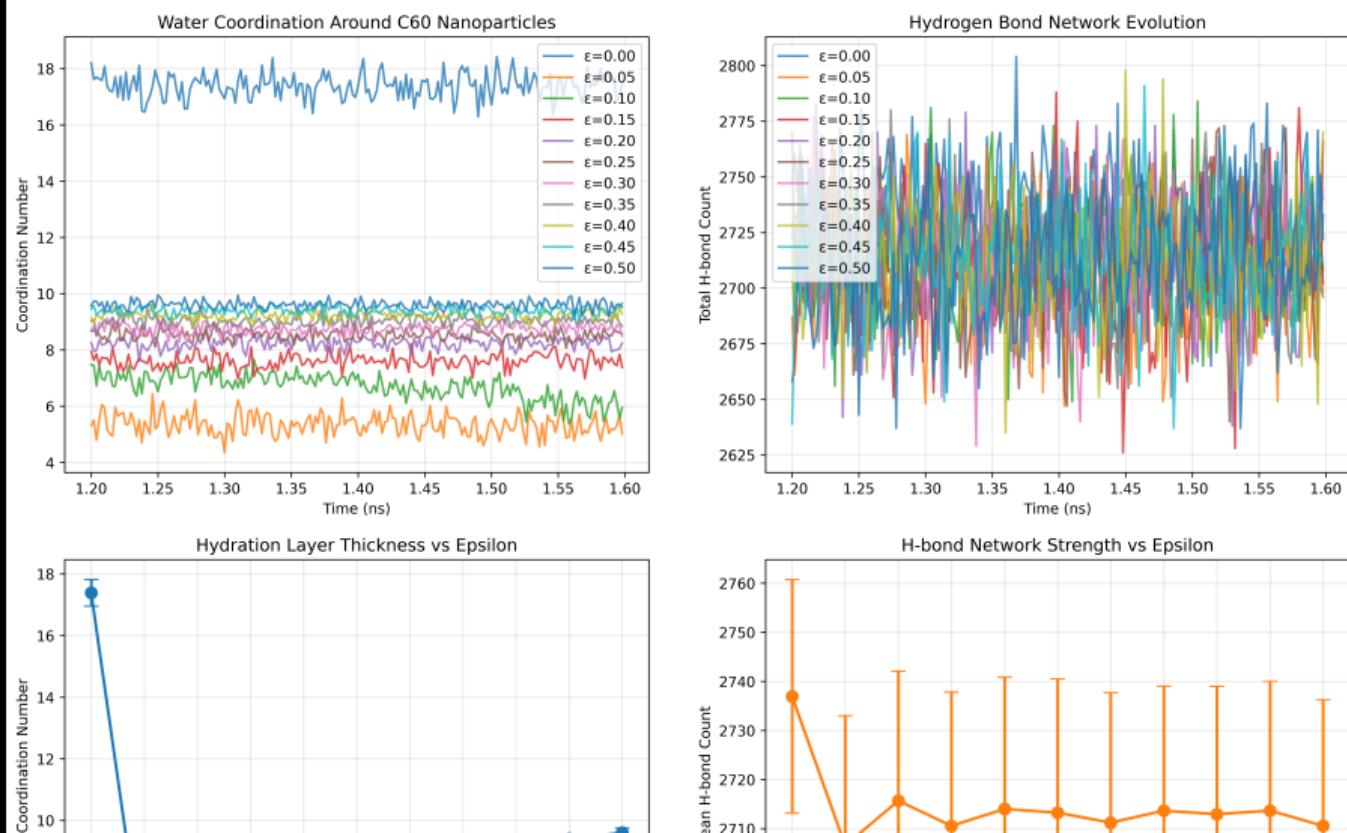
# 15 Shape Parameters (Asphericity)

Water Cluster Shape Parameters: Asphericity (Oblate) & Acylindricity (Prolate)

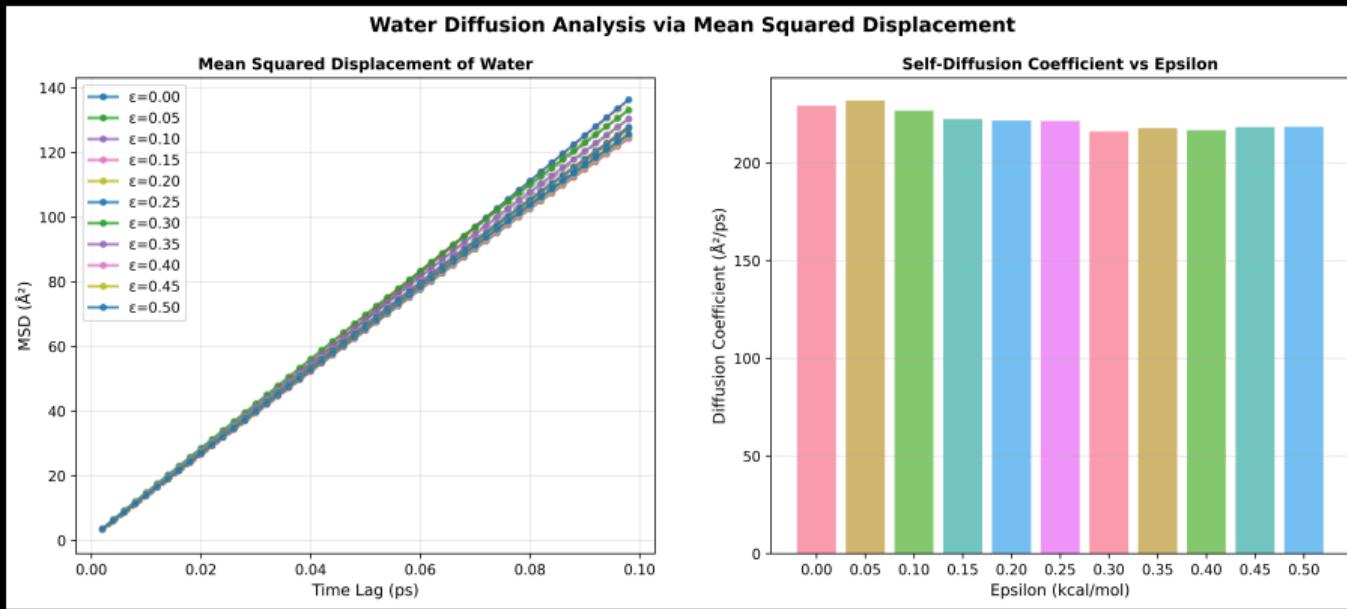


## 16 Coordination & H-bond Analysis

Water Coordination and Hydrogen Bonding Analysis



## 17 MSD / Diffusion Analysis



`extbfMSD:`  
`mathrmMSD(t)=`  
`langle —`  
`mathbfr(t)-`  
`mathbfr(0) —2`  
`rangle.`  
Diffusion:  $D =$   
 $\lim_t$   
to  
 $\infty$   
`mathrmMSD(t)/(6t).`

textwidth  
textwidth.

- ▶ Inter-fullerene distance time-series and aggregation metrics.

- ▶ Thermodynamic response functions via fluctuations:  $C_V =$

langle

delta  $E^2$

rangle/( $k_B T^2$ ),

κ<sub>T</sub> =

langle

delta  $V^2$

rangle/( $k_B T$

langle  $V$

rangle).

textwidth  
textwidth.

## 20 Aggregation Analysis

textwidth  
textwidth.

textwidth  
textwidth.

textwidth  
textwidth.

- ▶ Brightness: mean pixel intensity  
 $\mu_I$ ; Contrast:  
 $\sigma_I$  (std dev).

## 23 PPM Snapshot Montage

textwidth  
textwidth.

textwidth  
textwidth.

## Module 09: Equilibration Pathway - Math

extbfStages: NVT Thermalization, Pre-equilibration, Pressure ramp, NPT equilibration.  
Use multi-panel evolution plots to show stage-wise convergence.

textwidth  
textwidth.

- ▶ Bond length distributions, hydration shell counts (e.g.  $r_{C-O} < r_{shell}$ ).
- ▶ Displacement vectors

Delta

$\mathbf{m}_{\text{fr}} =$

$\mathbf{m}_{\text{fr}}^{\text{final}} - \mathbf{m}_{\text{fr}}^{\text{initial}}$ .

## 26 Bond Lengths

textwidth  
textwidth.

## 27 Hydration Shell

textwidth  
textwidth.

## 28 Radial Distributions (static)

textwidth  
textwidth.