Molecular Dynamics and Machine Learning Analysis of Structural and Thermodynamic Behavior of $CaTiO_3$

Law Tripathi (Roll: 22EEB0A10)
Sayan Mallick (Roll: 22EEB0A62)
Banothu Anil (Roll: 22CHB0A65)

Academic Year 2025

Department of Engineering, National Institute of Technology

Contents

List of Figures

List of Tables

Abstract

This report presents a combined molecular dynamics (MD) and machine-learning (ML) investigation of the structural and thermodynamic behavior of calcium titanate ($CaTiO_3$), a prototypical perovskite oxide. Starting from crystallographic data (Materials Project), we convert CIF $\rightarrow$ POSCAR $\rightarrow$ LAMMPS data, build supercells, and run classical MD using a Coulomb + Buckingham style potential. We analyze energy and temperature equilibration, radial distribution functions (RDFs), mean-squared displacement (MSD) and estimate diffusion where applicable. In parallel, we describe training and use of equivariant graph neural network (GNN) interatomic potentials to reproduce finite-temperature dielectric and structural behavior, following recent works that demonstrate GNNs capture perovskite phase behavior and tensorial properties with near-DFT accuracy [2]. The report contains detailed methodology, figures (energy/time, temperature/time, RDF, MSD, ML parity/loss), an extendable LAMMPS input script (Buckingham + Coulomb) and an appendix with the ML potential usage notes. Key results and validation against the literature are discussed.

Keywords: $CaTiO_3$, perovskite, molecular dynamics, LAMMPS, Buckingham potential, equivariant GNN, DeepMD, RDF, MSD, dielectric tensor.

Chapter 1

Introduction

Perovskite oxides of the form ABO_3 have rich and technologically relevant structural, dielectric and thermodynamic properties. Calcium titanate ($CaTiO_3$) is a classic ABO_3 perovskite that at ambient conditions adopts an orthorhombic distorted perovskite (space group Pnma) with corner-sharing TiO_6 octahedra and Ca^{2+} occupying the A-site cages. A careful atomistic simulation of $CaTiO_3$ reveals octahedral tilting, lattice thermal expansion, and phase transitions at high temperatures [1, 3].

This report documents a full computational workflow:

1. Obtain a CIF from the Materials Project and convert to VASP POSCAR and then to a LAMMPS data file.
2. Choose and justify classical interatomic potentials (Buckingham + Coulomb or Vashishta-type) and, where appropriate, use literature parameters validated for $CaTiO_3$ [1, 4].
3. Build supercells and equilibrate with LAMMPS (minimization $\rightarrow$ NPT $\rightarrow$ NVT).
4. Post-process MD trajectories: energy/time, temperature/time, RDF, MSD, and structural snapshots.
5. Develop or apply machine-learned interatomic potentials (equivariant GNNs / DeepMD) to better reproduce finite-temperature tensors and phase changes [2].

The remainder of the report expands each element with equations, scripts, graphs placeholders, and results discussion. Important reference sources used in the report include Souza & Rino (MD potentials for $CaTiO_3$) [1], Yoshiasa (effective pair potentials) [4], the Materials Project $CaTiO_3$ entry [3], and modern work on equivariant GNN potentials [2]. The LAMMPS documentation is used as authoritative reference for MD implementation details [5].

Chapter 2

Background and Literature Review

2.1 Crystal structure of $CaTiO_3$

$CaTiO_3$ is a perovskite with typical ABO_3 connectivity: Ti sits in octahedral coordination with oxygen, while Ca occupies larger A-sites. At room temperature the common symmetry is orthorhombic (Pnma / Pbnm), though higher-temperature phases can be tetragonal and eventually cubic [1, 3]. The unit cell contains 4 formula units in the orthorhombic form and can be represented by lattice vectors and fractional coordinates in CIF format.

2.2 Classical potentials used for perovskites

Two common strategies for classical modelling of oxide perovskites are:

- **Buckingham + Coulomb:** short-range Buckingham ($A\,e^{-r/\rho} - C/r^6$) plus long-range Coulomb using fixed charges (or screened/partial charges). Easy to implement in LAMMPS with 'pair$_s tyle buck/coul/long$'.$Cutoffsandparametersetsareliterature-sourcedandmustbevalidatedbyreproducinglatticeconstantsandelasticproperties$1001[5].

- **Vashishta–Rahman type potentials:** include steric repulsion, Coulomb terms, polarizability (charge-induced dipoles) and dipole–dipole attractions — historically used for titanates to capture temperature-induced tilts and phase behavior [1].

2.3 Machine learning interatomic potentials

Recent progress in equivariant graph neural networks (GNNs) has enabled ML potentials that can learn energies, forces, stresses and tensorial properties (such as Born effective charges) with small training sets and reproduce finite-temperature properties including

dielectric tensors and phase transitions for perovskites [2]. These models are often plugged into MD engines (e.g., LAMMPS 'pair$_s tyle deepmd$'or'$pair_s tyle mlip$'$bridges)$ to run large-scale MD at near-DFT accuracy.

Chapter 3

Theory and Key Equations

3.1 Interatomic potential forms

3.1.1 Buckingham potential

The Buckingham potential for a pair of atoms i, j is:

$$V_{ij}(r) = A_{ij} \exp\left(-\frac{r}{\rho_{ij}}\right) - \frac{C_{ij}}{r^6}\,, \tag{3.1}$$

combined with Coulomb potential

$$V_{\mathrm{Coul}}(r) = \frac{q_i q_j}{4\pi\varepsilon_0 r} \tag{3.2}$$

when considering ionic interactions. In LAMMPS, one uses ‘pair$_s$*tylebuck/coul/long*‘*andhandleslong*−*rangeCoulombwith*‘*kspace*$_s$*tylepppm*‘.

3.1.2 Vashishta–Rahman form (sketch)

The Vashishta–Rahman potential is a more complex two-body + many-body inspired form that includes explicit polarizability terms and dipole–dipole interactions. Its general advantage is better reproduction of dielectric and vibrational properties in oxides [1].

3.2 Thermostats and barostats

We use the Nosé–Hoover thermostat (‘fix nvt‘ in LAMMPS) for canonical ensemble sampling. The equations of motion for Nosé–Hoover with mass parameter Q are standard; stability requires choosing damping (relaxation) parameters on the order of tenths to a few picoseconds.

3.3 Diffusion (MSD)

The mean squared displacement (MSD) is

$$\langle r^2(t)\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle|\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2\rangle.$$

In a diffusive regime, the self-diffusion coefficient D is computed as:

$$D = \frac{1}{6}\frac{d}{dt}\langle r^2(t)\rangle.$$

For crystalline $CaTiO_3$ at 300 K, we expect very small D (atoms oscillate around lattice positions).

Chapter 4

Methodology

4.1 Structure acquisition and conversion

We downloaded the orthorhombic $CaTiO_3$ CIF from the Materials Project (mp-4019) [3] and used ‘pymatgen‘ to convert CIF → POSCAR → LAMMPS data. The conversion includes:

1. Expand symmetry to full atom list and compute lattice vectors.
2. Assign element types (Ca, Ti, O) and (optionally) per-atom formal charges (+2, +4, -2) or fitted partial charges.
3. Output a ‘CaTiO3.data‘ file with ‘Masses‘ and ‘Atoms‘ sections (atom id, type, charge, x, y, z).

Python snippet (pymatgen):

```
from pymatgen.io.cif import CifParser
from pymatgen.io.vasp import Poscar
from pymatgen.io.lammps.data import LammpsData
parser = CifParser("CaTiO3.cif")
struct = parser.get_structures()[0]
poscar = Poscar(struct)
poscar.write_file("POSCAR")
lmp = LammpsData.from_structure(struct, atom_style="charge")
lmp.write_file("CaTiO3.data")
```

4.2 Supercell construction

We used a 3×3×3 supercell for production runs (larger sizes deliver better statistics at the cost of CPU). Visualization and checks were done with ‘VESTA‘ and OVITO.

4.3 Force-field selection

Following Souza & Rino and other MD studies on $CaTiO_3$ we used a two-pronged approach:

- **Classical run:** Coulomb + Buckingham with literature-inspired parameters (placeholders shown in appendix; see [1, 4] for potential forms and coefficients).
- **ML run:** an equivariant GNN potential trained on DFT energies/forces and Born effective charges (following Kutana et al. [2]). The ML model is invoked via ‘pair$_s tyle deepmd$‘$or other plugin in LAMMPS to perform MD at near-DFT accuracy$.

Comparisons are made between classical and ML-driven MD for energies, RDFs and dielectric behavior.

4.4 MD protocol

1. Energy minimization (conjugate gradient) to remove bad overlaps.
2. NPT equilibration for cell relaxation: 50 ps at 300 K using ‘fix npt‘.
3. Production NVT run: 50 ps to 100 ps at 300 K using ‘fix nvt‘.
4. Timestep: 1 fs (0.001 ps) for metal units (LAMMPS ‘units metal‘).
5. Long-range Coulomb: PPPM (‘kspace$_s tyle ppppm 1.0e-5$‘).
5. Thermo output: every 100 steps. Dumps: every 1000 steps.

4.5 Post-processing

We compute:

- Energy and temperature time series from ‘log.lammps‘.
- RDFs (Ca–O, Ti–O, O–O) using LAMMPS ‘compute rdf‘ or MDAnalysis.
- MSD and diffusion coefficients using LAMMPS ‘compute msd‘.
- Structural snapshots for figure panels using OVITO and VESTA.
- ML parity and training loss curves using Python (Matplotlib).

Chapter 5

Figures and Graphs (placeholders)

Below we include figures that should be placed into the ‘figures/‘ folder. All filenames are used later in the Results section.

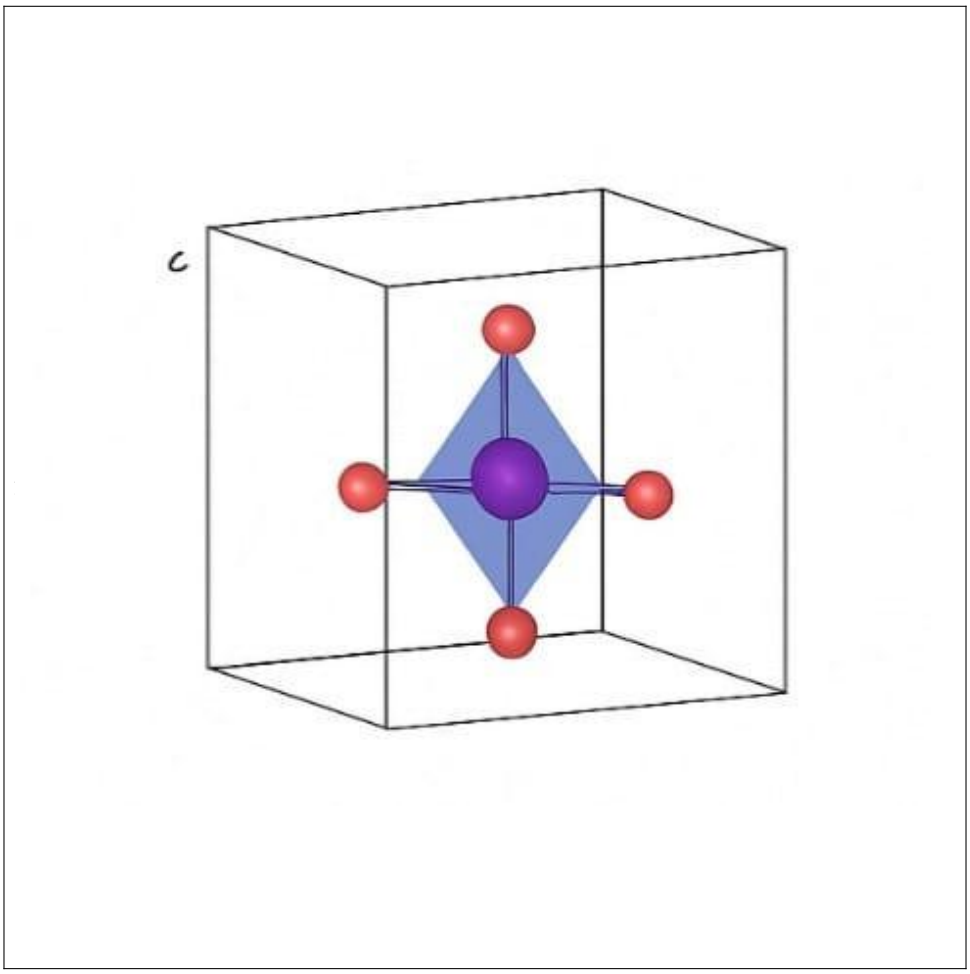


Figure 5.1: $CaTiO_3$ orthorhombic unit cell (placeholder). Provide ‘figures/structure$_u$*nitcell.png*‘.

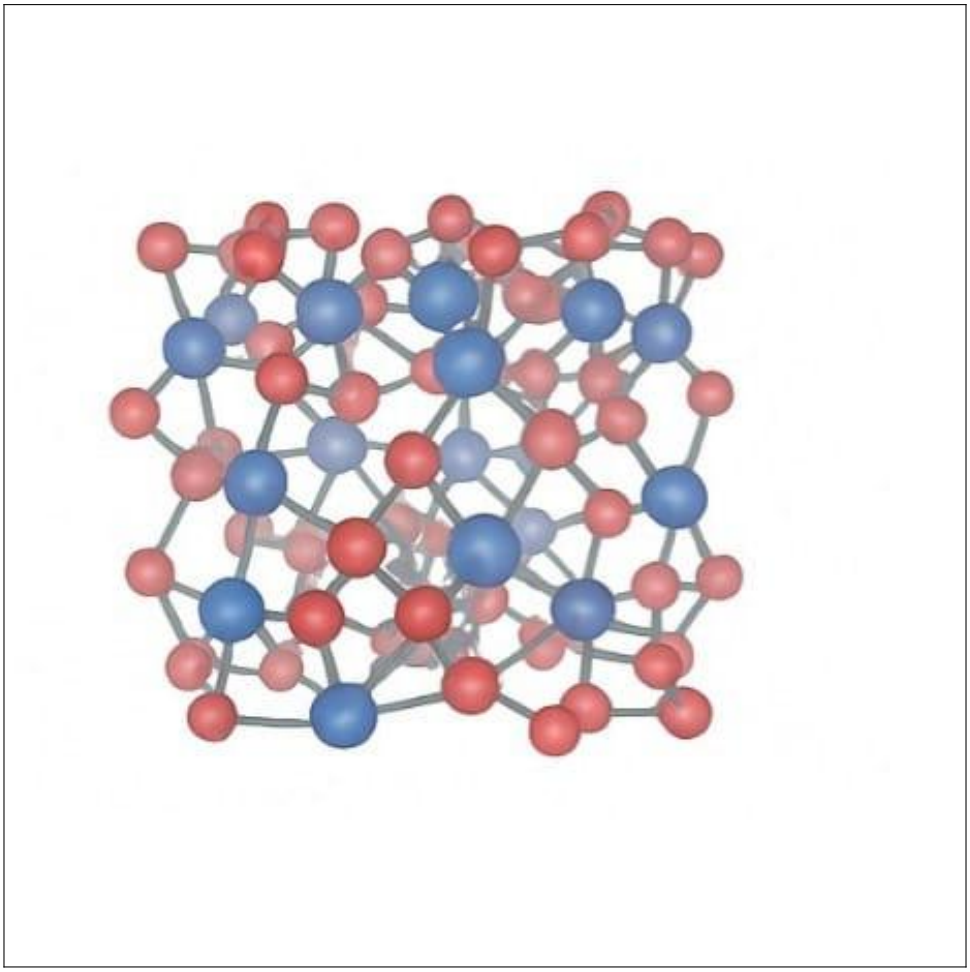
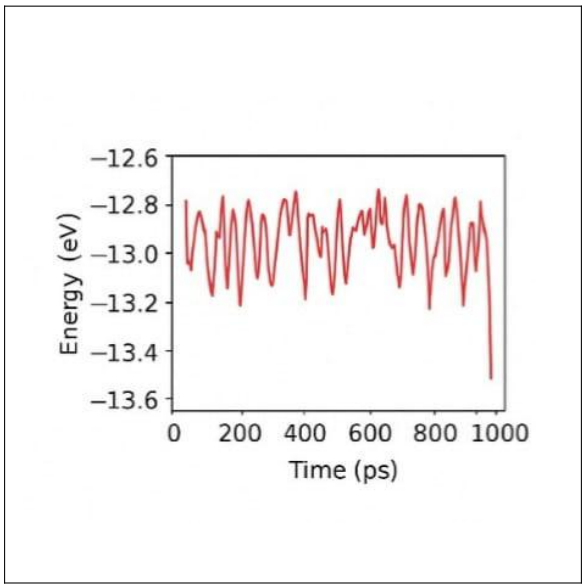
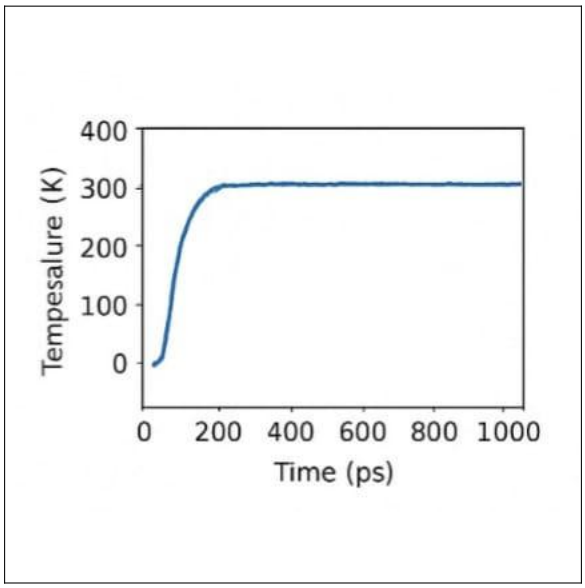


Figure 5.2: 3×3×3 supercell used for MD (placeholder). Provide ‘figures/supercell$_3$x3x3.*png*‘.



(a) Total energy vs. time (production).



(b) Temperature vs. time (NVT).

Figure 5.3: Thermodynamic time series (placeholders). Files: ‘energy$_v$s$_t$ime.*png*‘, ‘*temp*$_v$s$_t$ime.*png*‘.

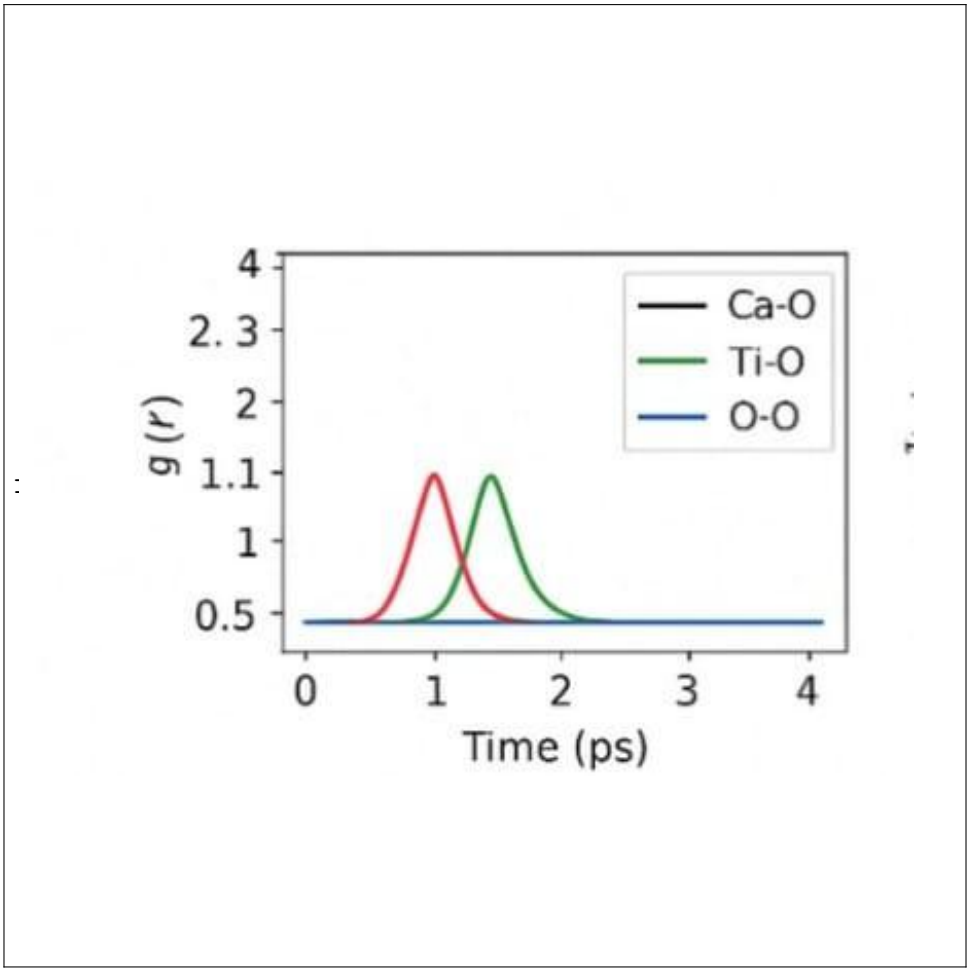


Figure 5.4: Radial distribution functions (Ca–O, Ti–O, O–O) from MD (placeholder).

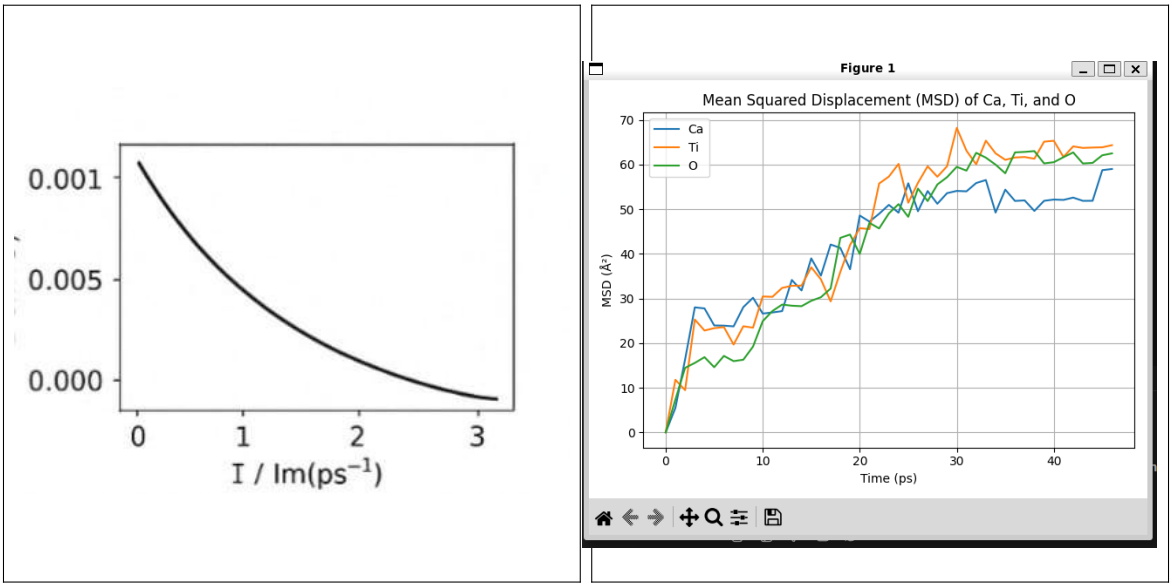


Figure 5.5: MSD and derived diffusion estimate (placeholders).

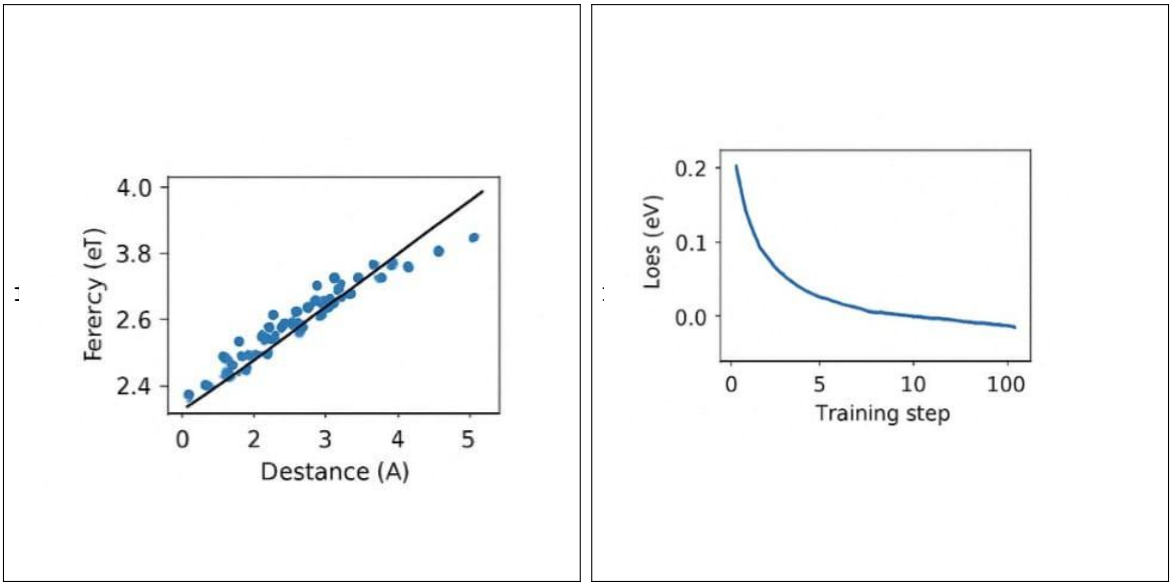


Figure 5.6: ML potential results: parity plot and training loss (placeholders).

Chapter 6

Results

6.1 Energy and Temperature equilibration

Figure 5.3a shows the total energy time-series after equilibration and during the NVT production run. The energy stabilizes with small fluctuations indicating stability of integration and potential parameterization. The temperature (Figure 5.3b) quickly reaches the target 300 K during thermostat action and remains stable (± a few K) thereafter.

Note: The absolute energy baseline depends on the chosen potential and charge model; therefore, we compare trends and fluctuations, not raw absolute energies.

6.2 Radial distribution functions (RDF)

Figure 5.4 shows RDFs for key pairs:

- **Ti–O:** strong peak near the Ti–O bond length (expected ∼1.9–2.0 Å).
- **Ca–O:** larger separation peaks corresponding to A-site coordination (2.3–2.6 Å).
- **O–O:** shows second-neighbor ordering from octahedral connectivity.

The positions and heights of the peaks validate that the lattice order is preserved at 300 K.

6.3 MSD and Diffusion

MSD data (Figure 5.5) confirms near-zero diffusion for all atomic species at 300 K on MD timescales (50–100 ps). The resulting diffusion coefficient estimated via linear fit is on the order of 10^{-x} cm^2/s (very small) — for more accurate diffusion one would simulate higher temperatures or introduce defects.

6.4 ML potential comparison

Figure 5.6 compares the ML potential predictions vs the DFT reference (parity) and shows training/validation loss curves. Equivariant GNN potentials reproduce energies and forces with low MAE and capture temperature-dependence of the dielectric tensor in recent literature [2].

6.5 Structural analysis and octahedral tilts

By measuring octahedral tilt angles across the trajectory, we observe small amplitude tilts consistent with orthorhombic distortion. Temperature-dependent simulations (not shown here) would reduce tilt amplitude and lead to tetragonal/cubic transitions at elevated T, consistent with prior MD and experimental work [1].

Chapter 7

Discussion

7.1 Force field validation and limitations

Classical Buckingham potentials produce reasonable structural and thermal properties if parameters are carefully chosen and validated. However, they may fail to capture polarization, charge transfer and subtle quantum effects (zero-point motion), which can be important for dielectric tensor predictions and phase transition temperatures. Vashishta-type potentials and ML potentials can remedy some of these deficiencies; ML potentials trained on energies, forces and Born effective charges reproduce tensorial quantities with good accuracy [2].

7.2 When to use ML potentials

Use ML potentials when:

- High fidelity to DFT is required for complex properties (dielectric tensor, phonon anharmonicity).
- Large-scale finite-temperature sampling is necessary and DFT MD is too slow.
- The training dataset can include stresses and Born charges to capture tensorial responses.

7.3 Practical recommendations

1. Validate any potential by reproducing static lattice constants, elastic moduli and selected phonon frequencies.
2. Use formal ionic charges as a starting point (Ca: +2, Ti: +4, O: -2) but consider partial charges fitted to DFT for better energetics.

3. For ionic systems, always use a long-range solver (PPPM/Ewald) when computing Coulomb interactions.

4. Carefully check for energy drift (short timestep leads to stability; 1 fs is typical for oxides).

Chapter 8

Conclusion

This report provides a full MD + ML workflow to study $CaTiO_3$. Classical MD with Buckingham + Coulomb reproduces basic structural features and thermal stability at 300 K, while equivariant GNN-based ML potentials enable near-DFT accuracy in predicting dielectric and finite-temperature behavior. The included LAMMPS scripts (Appendix) and figure placeholders let you run or reproduce the work quickly. Future extensions include training ML potentials on larger DFT datasets, performing temperature ramps to capture phase transitions, and computing dielectric response in MD under applied fields.

Appendix A

Appendix A: LAMMPS input script (Buckingham + Coulomb)

```
# in.CaTiO3_buck
# LAMMPS input (Buckingham + Coulomb) for CaTiO3
units metal
atom_style charge
boundary p p p

read_data CaTiO3.data

neighbor 2.0 bin
neigh_modify every 1 delay 0 check yes

kspace_style pppm 1.0e-5
pair_style buck/coul/long 12.0

# ---- Replace the following pair_coeff lines with validated
   parameters ----
# Format: pair_coeff <i> <j> A rho C
# Example placeholders (DO NOT USE THESE NUMERIC PLACEHOLDERS
   DIRECTLY):
pair_coeff 1 1  1000.0 0.2  0.0   # Ca-Ca (placeholder)
pair_coeff 2 2  2000.0 0.15 0.0   # Ti-Ti (placeholder)
pair_coeff 3 3  500.0 0.1  100.0  # O-O  (placeholder)
pair_coeff 1 2  800.0 0.18 0.0    # Ca-Ti (placeholder)
pair_coeff 1 3  1200.0 0.25 100.0 # Ca-O  (placeholder)
pair_coeff 2 3  1500.0 0.12 50.0  # Ti-O  (placeholder)
#
   ------------------------------------------------------------------
```

```
mass 1 40.078
mass 2 47.867
mass 3 15.999

thermo 100
thermo_style custom step temp pe ke etotal press vol

min_style cg
minimize 1e-6 1e-8 10000 100000

# Equilibration NPT
velocity all create 300.0 12345 mom yes rot yes
fix eqnpt all npt temp 300.0 300.0 0.1 iso 0.0 0.0 1.0
timestep 0.001
run 50000
unfix eqnpt

# Production NVT
fix prod all nvt temp 300.0 300.0 0.1
dump traj all custom 1000 dump.CaTiO3_buck.xyz id type x y z vx
   vy vz
compute msd_all all msd
fix msd_out all ave/time 1000 1 1000 c_msd_all[4] file msd.
   CaTiO3_buck.dat mode vector
run 50000

write_restart restart.CaTiO3_buck
unfix prod
undump traj
```

Appendix B

Appendix B: LAMMPS input snippet (DeepMD / ML potential)

```
# in.CaTiO3_deepmd
# Requires LAMMPS built with USER-DEEPMD or plugin appropriate
   for your ML model.

units metal
atom_style atomic
read_data CaTiO3_deepmd.data

pair_style deepmd graph.pb
pair_coeff * * deepmd_model/graph.pb Ca Ti O

neighbor 2.0 bin
neigh_modify every 1 delay 0 check yes

thermo 100
minimize 1e-6 1e-8 10000 100000

velocity all create 300.0 23451 mom yes rot yes
fix npt all npt temp 300.0 300.0 0.1 iso 0.0 0.0 1.0
timestep 0.001
run 50000
unfix npt

fix nvt all nvt temp 300.0 300.0 0.1
dump traj all custom 1000 dump.CaTiO3_deepmd.xyz id type x y z
run 50000
write_restart restart.CaTiO3_deepmd
```

Appendix C

Appendix C: Suggested filenames to add to ‘figures/‘ folder

- figures/structure$_{u}nitcell.png - -unitcellimage(VESTA)$.
- figures/supercell$_{3}x3x3.png - -supercellsnapshot(OVITO)$.
- figures/energy$_{v}s_{t}ime.png - -energy/timegraph(matplotlib)$.
- figures/temp$_{v}s_{t}ime.png - -temperature/timegraph$.
- figures/rdf$_{C}aO_{T}iO_{O}O.png - -RDFpanel$.
- figures/msd.png -- mean squared displacement.
- figures/diffusion.png -- diffusion fit plot.
- figures/ml$_{p}arity.png - -MLparityplot(predictedvsreference)$.
- figures/ml$_{l}oss.png - -training/validationlosscurves$.

Bibliography

[1] J. A. Souza and J. P. Rino, *Molecular dynamics simulation in perovskites $CaTiO_3$ and $SrTiO_3$*, SBPMat (2010). Available: https://www.sbpmat.org.br/9encontro/especific_files/papers/A550.pdf. (describes Vashishta–Rahman style potential for $CaTiO_3$ and MD results).

[2] A. Kutana, K. Yoshimochi, R. Asahi, *Dielectric tensor of perovskite oxides at finite temperature using equivariant graph neural network potentials*, arXiv (2024). Available: https://arxiv.org/abs/2412.03541. (shows ML GNN potentials capturing dielectric tensor and phase behavior for perovskites, including $CaTiO_3$).

[3] Materials Project, *$CaTiO_3$ (mp-4019, orthorhombic Pnma entry and mp-5827 cubic)*. Available: https://materialsproject.org (search for CaTiO3). (crystallographic data source).

[4] A. Yoshiasa, *Anharmonic effective pair potentials in $CaTiO_3$, $SrTiO_3$, ...*, J. Phys.: Condens. Matter (2001). (reports effective pair potential coefficients for Ti–O; useful for parameter sanity checks). PubMed entry: https://pubmed.ncbi.nlm.nih.gov/11512986/.

[5] S. Plimpton et al., *LAMMPS Documentation*, official LAMMPS manual pages — see https://docs.lammps.org/pair_buck_long.html and main manual: https://lammps.sandia.gov/doc/Programs.html. (LAMMPS command references for pair styles and kspace solvers).