Analysis of MD Results Using Statistical Mechanics Methods

Ioan Kosztin
Beckman Institute
University of Illinois at Urbana-Champaign

Molecular Modeling

1. Model building
2. Molecular Dynamics Simulation
3. Analysis of the
   • model
   • results of the simulation
Collection of MD Data

- DCD trajectory file
  - coordinates for each atom
  - velocities for each atom

- Output file
  - global energies
  - temperature, pressure, …

Analysis of MD Data

1. Structural properties
2. Equilibrium properties
3. Non-equilibrium properties

Can be studied via both equilibrium and non-equilibrium MD simulations
Equilibrium (Thermodynamic) Properties

MD simulation → microscopic information → macroscopic properties

Statistical Mechanics

\[ \Gamma[r(t), p(t)] \]

Phase space trajectory

\[ \rho(\Gamma) \]

Ensemble average over probability density

Statistical Ensemble

Collection of large number of replicas (on a macroscopic level) of the system

Each replica is characterized by the same macroscopic parameters (e.g., NVT, NPT)

The microscopic state of each replica (at a given time) is determined by \( \Gamma \) in phase space
**Time vs Ensemble Average**

For $t \to \infty$, $\Gamma(t)$ generates an ensemble with

$$\rho(\Gamma)d\Gamma = \lim_{t \to \infty} d\tau / t$$

**Ergodic Hypothesis:** Time and Ensemble averages are equivalent, i.e.,

$$\langle A(r, p) \rangle_t = \langle A(\Gamma) \rangle_\rho$$

**Time average:**

$$\langle A \rangle_t = \frac{1}{T} \int_0^T dt \, A[r(t), p(t)]$$

**Ensemble average:**

$$\langle A \rangle = \int d\Gamma \, \rho(\Gamma) \, A(\Gamma)$$

---

**Thermodynamic Properties from MD Simulations**

Thermodynamic (equilibrium) averages can be calculated via time averaging of MD simulation time series

$$\langle A \rangle \approx \frac{1}{N} \sum_{i=1}^{N} A(t_i)$$

Finite simulation time means incomplete sampling!
Common Statistical Ensembles

1. Microcanonical \((N,V,E)\):
   \[ \rho_{NVE}(\Gamma) \propto \delta[H(\Gamma) - E] \] ← Newton’s eq. of motion

2. Canonical \((N,V,T)\):
   \[ \rho_{NVT}(\Gamma) = \exp\left\{\left[F - H(\Gamma)\right]/k_B T\right\} \] ← Langevin dynamics

3. Isothermal-isobaric \((N,p,T)\)
   \[ \rho_{NPT}(\Gamma) = \exp\left\{\left[G - H(\Gamma)\right]/k_B T\right\} \] ← Nose-Hoover method

Different simulation protocols \([\Gamma(t) \rightarrow \Gamma(t+\delta t)]\) sample different statistical ensembles

Examples of Thermodynamic Observables

• Energies (kinetic, potential, internal,...)
• Temperature [equipartition theorem]
• Pressure [virial theorem]

Thermodynamic derivatives are related to mean square fluctuations of thermodynamic quantities

• Specific heat capacity \(C_v\) and \(C_P\)
• Thermal expansion coefficient \(\alpha_p\)
• Isothermal compressibility \(\beta_T\)
• Thermal pressure coefficient \(\gamma_V\)
Mean Energies

Total (internal) energy: \[ E = \frac{1}{N} \sum_{i=1}^{N} E(t_i) \]

Kinetic energy: \[ K = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{M} \left( \frac{p_j^2(t_i)}{2m_j} \right) \]

Potential energy: \[ U = E - K \]

Note: You can conveniently use \texttt{namdplot} to graph the time evolution of different energy terms (as well as \( T, P, V \)) during simulation.

Temperature

From the equipartition theorem \( \langle p_k \frac{\partial H}{\partial p_k} \rangle = k_B T \)

\[ T = \frac{2}{3Nk_B} \langle K \rangle \]

Instantaneous kinetic temperature

\[ T = \frac{2K}{3Nk_B} \]

Note: in the NVTP ensemble \( N \rightarrow N-N_c \), with \( N_c = 3 \)
Pressure

From the virial theorem \( \langle r_k \partial H / \partial r_k \rangle = k_B T \)

\[
P V = N k_B T + \langle W \rangle
\]

The *virial* is defined as

\[
W = \frac{1}{3} \sum_{j=1}^{M} r_j \cdot f_j = - \frac{1}{3} \sum_{i,j>i} w(r_{ij})
\]

with \( w(r) = r \frac{d \nu(r)}{dr} \)

Instantaneous *pressure* function (not unique!)

\[
\tilde{P} = \rho k_B T + W / V
\]

Thermodynamic Fluctuations (TF)

\[
\langle \delta A \rangle \approx \frac{1}{N} \sum_{i=1}^{N} [A(t_i) - \langle A \rangle]
\]

Mean Square Fluctuations (MSF)

\[
\langle \delta A^2 \rangle = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2
\]

According to *Statistical Mechanics*, the probability distribution of thermodynamic fluctuations is

\[
\rho_{\text{fluct}} \propto \exp \left( \frac{\delta P \cdot \delta V - \delta T \cdot \delta S}{2k_B T} \right)
\]
TF in NVT Ensemble

In MD simulations distinction must be made between properly defined mechanical quantities (e.g., energy $E$, kinetic temperature $T$, instantaneous pressure $P$) and thermodynamic quantities, e.g., $T$, $P$, ...

For example: \[
\langle \delta E^2 \rangle = \langle \delta \mathcal{H}^2 \rangle = k_B T^2 C_V \quad \checkmark
\]

But: \[
\langle \delta P^2 \rangle \neq \langle \delta \mathcal{P}^2 \rangle = k_B T / V \beta_T \quad \times
\]

Other useful formulas:
\[
\langle \delta K^2 \rangle = \frac{3N}{2} (k_B T)^2
\]
\[
\langle \delta U^2 \rangle = k_B T^2 (C_V - 3Nk_B / 2)
\]
\[
\langle \delta U \delta \mathcal{P} \rangle = k_B T^2 (\gamma_V - \rho k_B)
\]

$C_V = (\partial E / \partial T)_V$

$\gamma_V = (\partial P / \partial T)_V$

TF in NPT Ensemble

\[
\langle \delta V^2 \rangle = V k_B T \beta_T
\]
\[
\langle \delta (\mathcal{H} + PV)^2 \rangle = k_B T^2 C_P
\]
\[
\langle \delta V \delta (\mathcal{H} + PV) \rangle = k_B T^2 V \alpha_P
\]

By definition: \[
\alpha_P = V^{-1} (\partial V / \partial T)_P ; \quad \beta_T = -V^{-1} (\partial V / \partial P)_T
\]
\[
C_P = (\partial E / \partial T)_P
\]
How to Calculate $C_V$?

1. From definition

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V$$

Perform multiple simulations to determine $E \equiv \langle E \rangle$ as a function of $T$,
then calculate the derivative of $E(T)$ with respect to $T$.

2. From the MSF of the total energy $E$

$$C_V = \frac{\langle \delta E^2 \rangle}{k_B T^2}$$

with

$$\langle \delta E^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2$$