Treatment of the nonbonded energy terms

Mamta Sagar
UIET, CSJMU Kanpur

- The most time consuming part of a molecular dynamics simulation is the calculation of the nonbonded terms in the potential energy function, e.g., the electrostatic and van der Waals forces.

$$\sum_{i,j}\left\{4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]+\frac{q_i q_j}{\varepsilon D r_{ij}}\right\}$$

- In principle, the non-bonded energy terms between every pair of atoms should be evaluated; in this case, the number of increases as the square of the number of atoms for a pairwise model (N2).
- To speed up the computation, the interactions between two atoms separated by a distance greater than a pre-defined distance, the cutoff distance, are ignored. Several different ways to terminate the interaction between two atoms have been developed over the years; some work better than others

- Truncation: the interactions are simply set to zero for interatomic distances greater than the cutoff distance. This method can lead to large fluctuations in the energy. This method is not often used.

- The SHIFT cutoff method: this method modifies the entire potential energy surface such that at the cutoff distance the interaction potential is zero. The drawback of this method is that equilibrium distances are slightly decreased.

- The SWITCH cutoff method: This method tapers the interaction potential over a predefined range of distances. The potential takes its usual value up to the first cutoff and is then switched to zero between the first and last cutoff. This model suffers from strong forces in the switching region which can slightly perturb the equilibrium structure. The SWITCH function is not recommended when using short cutoff regions.

Long range electrostatic interactions

- Inclusion of the longer range electrostatic interactions in a molecular dynamics simulation by simply increasing the cutoff distance can dramatically raise the computational cost. Most often, the long-range electrostatic interactions are ignored, however, in some cases, their neglect introduces a severe approximation; for example in the calculations of dielectric properties (Alper and Levy 1989) or in metal ion - protein interactions. (Stote and Karplus 1995).

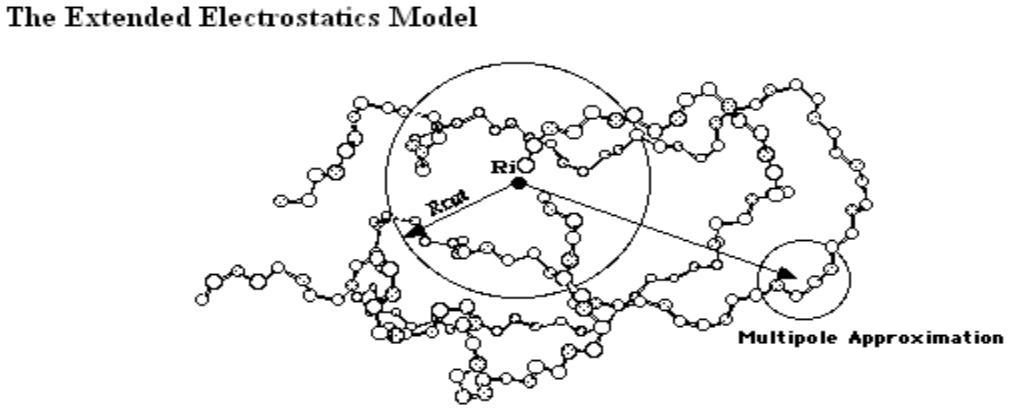
- In recent years, a number of models have been introduced which permit the inclusion of long-range electrostatic interactions in molecular dynamics simulation. For simulations of proteins and enzymes in a crystalline state, the Ewald summation is considered to be the correct treatment for long range electrostatic interactions (Allen and Tildesley 1989). Variations of the Ewald method for periodic systems include the particle-mesh Ewald method (York, Darden et al. 1993).

- To treat non-periodic systems, such as an enzyme in solution methods based on multipole expansions have been developed. Many of these methods partition the electrostatic interaction into a long-range component and a short range component.
- The short range component is treated in the usual pairwise fashion while a multipole approximation is introduced to approximate the long-range electrostatic interaction; several such models have been developed (Brooks, Bruccoleri et al. 1983; Stote, States et al. 1991; Shimada, Kaneko et al. 1994).

- Two such models have been implemented in the CHARMM program, the Extended Electrostatics model (Stote, States et al. 1991) and the Fast Multipole Method Method (Greengard and Rokhlin 1987; Shimada, Kaneko et al. 1994).
- Just to cite a couple of examples where these methods significantly improved the simulation, the Extended Electrostatics methods was used in a study of the binding interactions in the RNase A/3'-UMP enzyme-product complex (Straub, Lim et al. 1994).

The Extended Electrostatics Model

- The Extended Electrostatics model approximates the full electrostatic interaction by partitioning the electric potential and the resulting forces on the atom at Ri into a "Near" and an "Extended" contribution. The "Near" contribution arises from the charged particles which fall within the sphere defined by the cutoff distance Rcut, while the "Extended" contribution, arises from the particles which are beyond the cutoff distance Rcut. The "Near" contribution is calculated by a conventional pairwise sum and the "Extended" contribution to the potential at Ri is calculated using a multipole approximation (Stote, States et al. 1991).



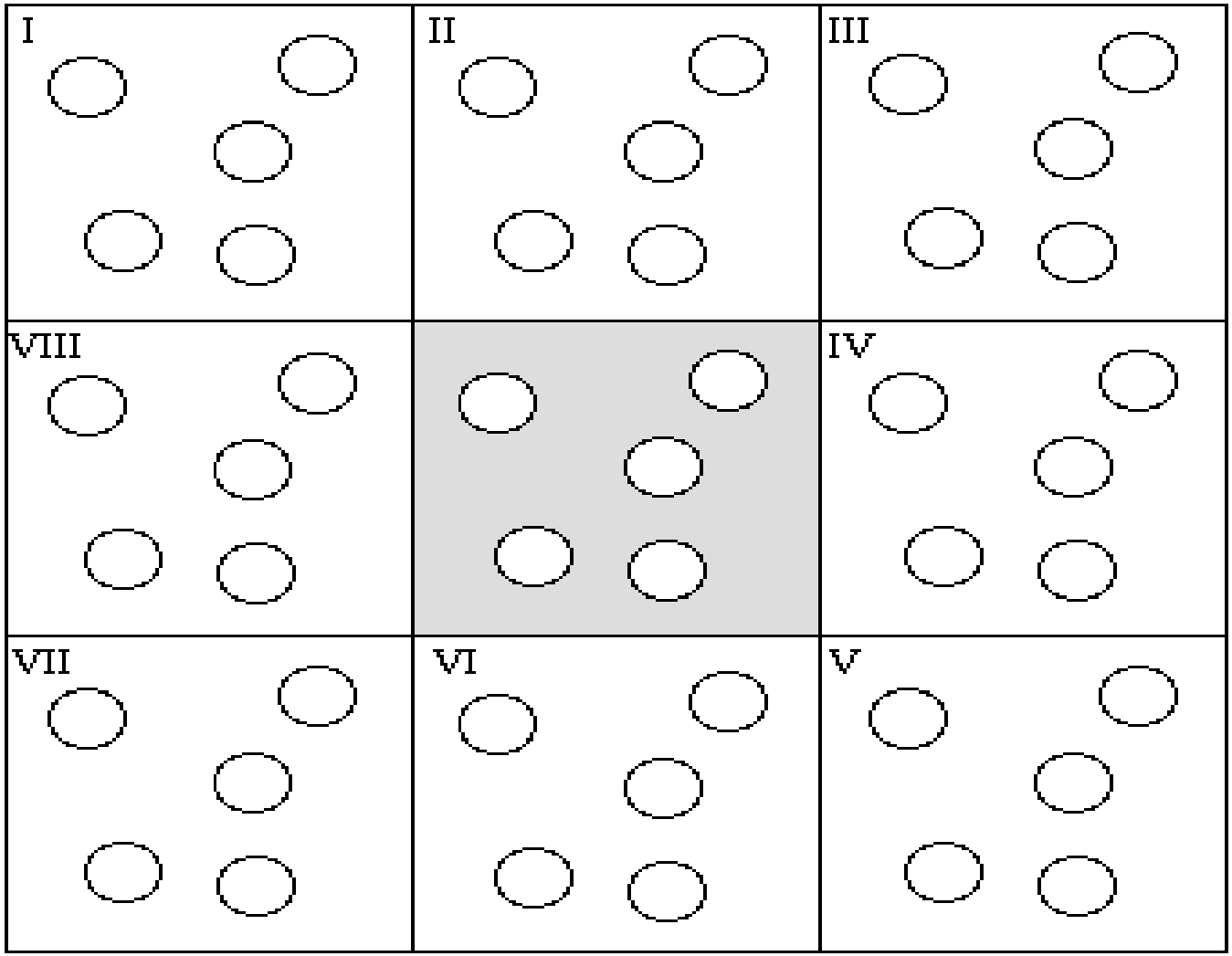
Treatment of Solvent in a Molecular Dynamics Simulation

- Solvent, usually water, has a fundamental influence on the structure, dynamics and thermodynamics of biological molecules, both locally and globally. One of the most important effects of the solvent is the screening of electrostatic interactions. The electrostatic interaction between two charges is given by Coulomb's law,

$$V_{elec} = \frac{q_i q_j}{\varepsilon_{eff} r_{ij}}$$

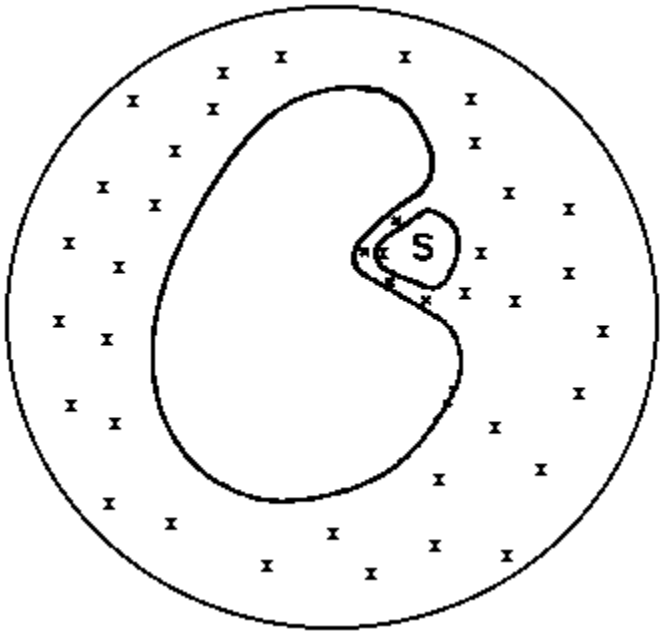
where q_i, q_j are the partial atomic charges, ε_{elec} is the effective dielectric constant and r_{ij} is the relative distance between the two particles. It is important to include solvent effects in an MD simulation. This can be done at several levels. The simplest treatment is to simply include a dielectric screening constant in the electrostatic term of the potential energy function. In this **implicit** treatment of the solvent, water molecules are not included in the simulation but an effective dielectric constant is used. Often the effective dielectric constant is taken to be distance dependent, $\varepsilon_{eff} = r_{ij}\varepsilon$), where ε ranges from 4 to 20. Although this is a crude approximation, it is still much better than using unscreened partial charges.

Periodic Boundary conditions



- Periodic boundary conditions enable a simulation to be performed using a relatively small number of particles in such a way that the particles experience forces as though they were in a bulk solution. See, for example, the two dimensional box.
- The central box is surrounded by eight neighbors. The coordinates of the image particles, those found in the surrounding box are related to those in the primary box by simple translations.
- The simplest box is the cubic box. Forces on the primary particles are calculated from particles within the same box as well as in the image box. The cutoff is chosen such that a particle in the primary box does not see its image in the surrounding boxes.

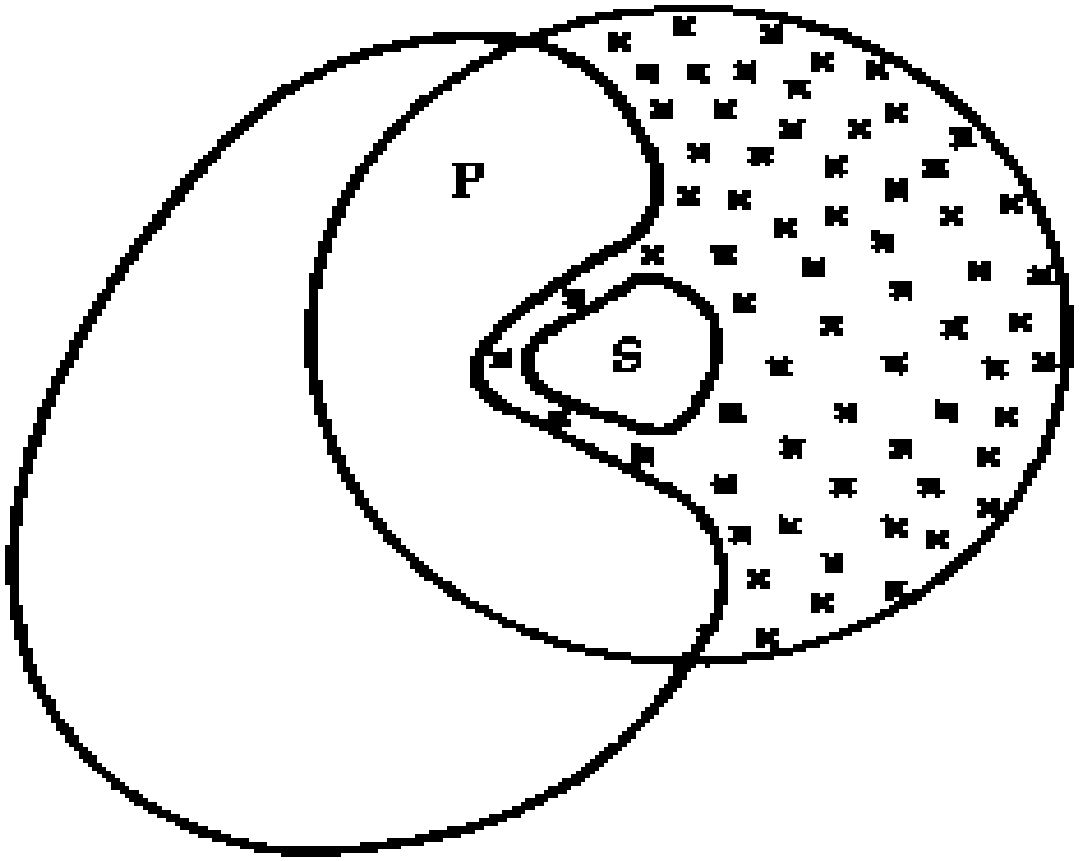
Solvation shells



There exist numerous cases where one may not wish to use periodic boundary conditions. In some cases, the use of periodic boundary conditions requires the use of a prohibitively large number of water molecules

- With the increase in computer power, it has become much more feasible to incorporate water molecules in the simulation. The simplest way is to surround the protein or just a part of the protein with a sphere of water. Boundary potentials have been developed which restrain the water molecules to a sphere while maintaining a strong resemblence to bulk water. Structural and thermodynamics properties when calculated under these conditions indicate that the water still behaves as bulk water. This usually involves much fewer water molecules than in a periodic boundary simulation and is often sufficient

Active site solvation

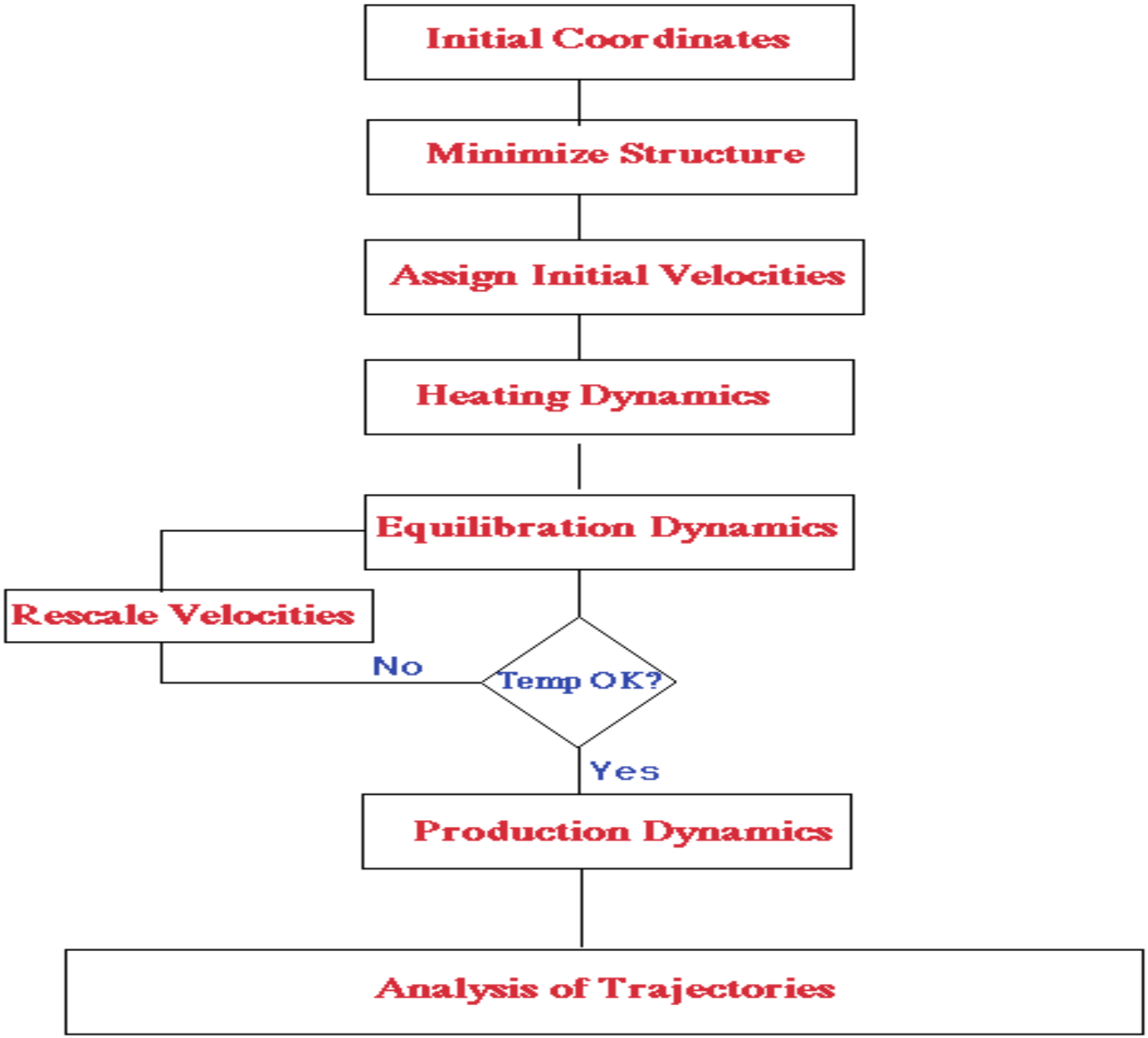


- Often in the case of proteins, in particular enzymes, there is a large protein scafold yet one is primarily interested in what is happening in the active site. In this case, the enzyme can be partitioned into several regions. The reaction zone corresponds to that part of the enzyme which is of interest, usually the active site. Everything outside the reaction zone is referred to as the reservoir region. Atoms in the reservoir region are usually held fixed or harmonically constrained. The reaction zone is then solvated with a sufficiently large sphere of water and only this region is allowed to move during a molecular dynamics simulation. This allows for a significant speed up of computer time if one is just interested in a localized region.

Setting up and running a Molecular Dynamics Simulations

- In a molecular dynamics simulation, the time dependent behavior of the molecular system is obtained by integrating Newton's equations of motion using one of the numerical integrators described earlier (see Classical Mechanics) and the potential energy function (see Potential Energy Function).
- The result of the simulation is a time series of conformations; this is called a trajectory or the path followed by each atom in accordance with Newton's laws of motion.

- Most molecular dynamics simulations are performed under conditions of constant N,V,E (the microcanonical ensemble), but more recent methods perform simulations at constant N, T and P to better mimic experimental conditions.


Initial Coordinates
Minimize Structure
Assign Initial Velocities
Heating Dynamics
Equilibration Dynamics
Rescale Velocities
No
Temp OK?
Yes
Production Dynamics
Analysis of Trajectories

I)Mean Energy

$$\langle E \rangle = \frac{1}{N} \sum_{i=1}^{N} E_i$$

II)RMS difference between two structures

$$RMS = \left\langle \left(r_i^{\alpha} - r_i^{\beta} \right)^2 \right\rangle^{\frac{1}{2}} = \sqrt{\frac{1}{N_i} \sum_i \left(r_i^{\alpha} - r_i^{\beta} \right)^2}$$

II) RMS difference between two structures

$$RMS = \left\langle \left(r_i^{\alpha} - r_i^{\beta} \right)^2 \right\rangle^{\frac{1}{2}} = \sqrt{\frac{1}{N_i} \sum_i \left(r_i^{\alpha} - r_i^{\beta} \right)^2}$$

III) RMS fluctuations

$$RMS_i^{fluct} = \sqrt{\frac{1}{N_f} \sum_f \left(r_i^f - r_i^{ave} \right)^2}$$

note the relation between the RMS fluctuations and the crystallographic B factors;

$$B_i = \frac{8}{3} \pi^2 \left(RMS_i^{fluct} \right)^2$$

IV) radius of gyration

$$RadiusGyration = \sqrt{\frac{1}{Ni} \sum_i \left(r_i - r_{cm} \right)^2}$$

where $r_i - r_{cm}$ is the distance between atom i and the center of mass of the molecule.

- From a molecular dynamics simulation, time dependent properties such as correlation functions can also be calculated. These, in turn, can be related to spectroscopic measurement.