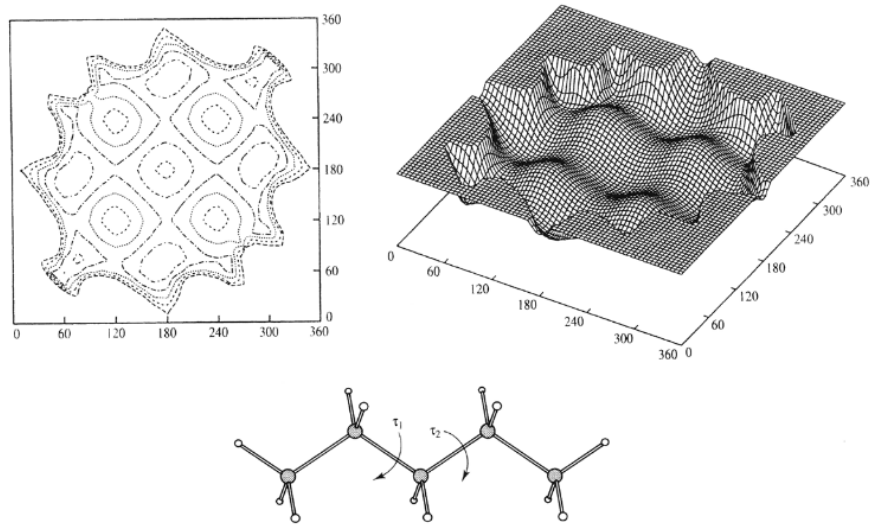
6.1 Energy Minimization

Introduction



Energetic landscape: The collection of points over the different variables $(Var1, Var2, ..., VarN, Energy)$ constitutes an Hyperplane.

Local Minima in this hyperplane correspond to energetically favorable conformations.

The global minimum corresponds to the fold of the molecule.

Ensemble average

Properties of molecular systems defined by the average over all molecules.

Ensemble average weighted by relevance: - Energetically unfavorable states do not contribute much. - Energetically favorable, more frequent, contribute much.

Ergodic hypothesis: Time average = ensemble average.

Strategies for searching minima

Local minima are easier to find than global minima.

Gradient descent methods

$$gradient = \nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, ..., \frac{\partial}{\partial x_n} \right)$$

The gradient points towards the steepest slope of f at a given point.

It's intuitive to follow this gradient until we reach a point with $\nabla = 0$ (or a threshold for the root mean square of the gradient previously defined), which corresponds to a minima, maxima or a saddle.

For k = 0 to convergence

- Choose direction $\mathbf{s}_k$
- Choose step size λ_k along $\mathbf{s}_k$, such that energy is minimized
- Update coordinates:

$$\mathbf{r}_{k+1} = \mathbf{r}_k + \lambda_k \mathbf{s}_k$$

- $k = k + 1$, next iteration

Steepest gradient descent (SD)

- Greedy algorithm
- Always chooses direction of steepest descent
- Continue until no further descent

For $k = 0$ to convergence

- calculate gradient $\mathbf{g}_k$
- $\mathbf{s}_k = -\mathbf{g}_k / |\mathbf{g}_k|$
- $\lambda_k = \textbf{new step}(\mathbf{r}_k, \mathbf{s}_k)$
- Update coordinates

$$\mathbf{r}_{k+1} = \mathbf{r}_k + \lambda_k \mathbf{s}_k$$

- If converged, stop, else
- $k = k + 1$

- Determining size of the step is critical.

Advantages:

- Simple, fast for steep gradients
- Numerically stable
- Fast and accurate for convex functions

Disadvantages: - Slow convergence - Search directions can be orthogonal - Can result in zig zag searches

Conjugated gradient descent

It keeps memory of the last search direction. $s = -g_k + \beta_{k+1} s_{k-1}$

Where β is the weight of the last search direction, a parameter than needs to be set properly.

Conformational search methods

Gradient descent methods are able to find only local minima, which can correspond to the global maximum or not.

Other methods need to be used if we want to find the global minimum for sure.

Systematic (grid based) search

Only feasible for a small number of degrees of freedom. It leads to combinatorial explosion very fast.

The number of degrees of freedom can be reduced drastically: - Coordinates are not independent - Flexibility lies in torsions - Not all dihedral angles are feasible - The number of possible conformations is nonetheless too big.

Stochastic methods

Ensemble corresponds to phase space (6N) integral.

Random sampling can be used for estimating this integral → Monte-Carlo method

MonteCarlo Method

Randomly sample conformational space.

- For k iterations
  - Choose random point ($\mathbf{r}_k$, $\mathbf{p}_k$) in phase space
  - Determine properties of interest, $A_k(\mathbf{r}_k, \mathbf{p}_k)$
  - Calculate energy $E_k(\mathbf{r}_k, \mathbf{p}_k)$
  - Compute Boltzmann-Factor $\exp(-E_k(\mathbf{r}_k, \mathbf{p}_k)/(k_B T))$
  - Sum up weighted values of A
  - Sum up Boltzmann-Factors (to calculate the partition function)
- Compute ensemble mean as:

$$< A >= \sum_k A_k(\mathbf{r}, \mathbf{p}) \rho_k(\mathbf{p}, \mathbf{r})$$

Problem: Too many irrelevant conformations are sampled.

MonteCarlo method with importance sampling (MCMC)

The boltzmann distribution is used for building a markov chain with a sample rejection probability. $\frac{p_i}{p_j} = e^{-\frac{E_i - E_j}{k_B T}}$

A markov chain is a stochastic model, where the current state of the system only depends on the previous state (it has no memory).

- Für k Iterations
  - Randomly choose $\mathbf{r}_k$ in conformational space
  - If $E_k < E_{k\text{-}1}$
    - Accept new conformation
  - Otherwise, $E_k > E_{k\text{-}1}$
    - Draw a equally distributed random number, x [0, 1]
    - If $x < \exp(-(E_k - E_{k\text{-}1})/(k_B T))$
      - Accept new conformation
    - Otherwise
      - Retain old conformation

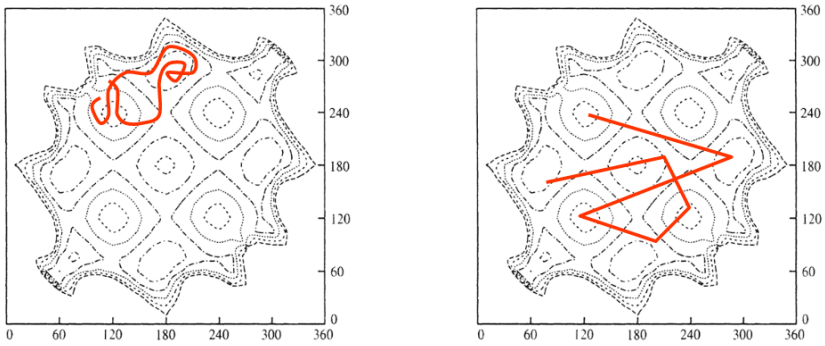
with probability of moving to a new point:

$$p(x_n, x_{n+1}) = min\left(\frac{f(x_n+1)}{f(x_n)}, 1\right)$$

So, if the value of the new point is bigger, it always moves to it, if not, there-s a rejection probability.

This simulation results in a trajectory and associated energy, but there's no temporal correspondence between succesive points in the trajectory. It's used to estimate average properties, but also to find minima as the trajectory samples favorable states more often.

- MD suitable for estimating time-dependent parameters (e.g.. diffusion coefficient)
- MC: faster convergence when computing thermodynamic properties
- MD: explores local conformational spatial neighborhood better
- MC: better explores global conformational space
- MD and MC are complementing each other
- combining both simulation types possible



Simulated annealing

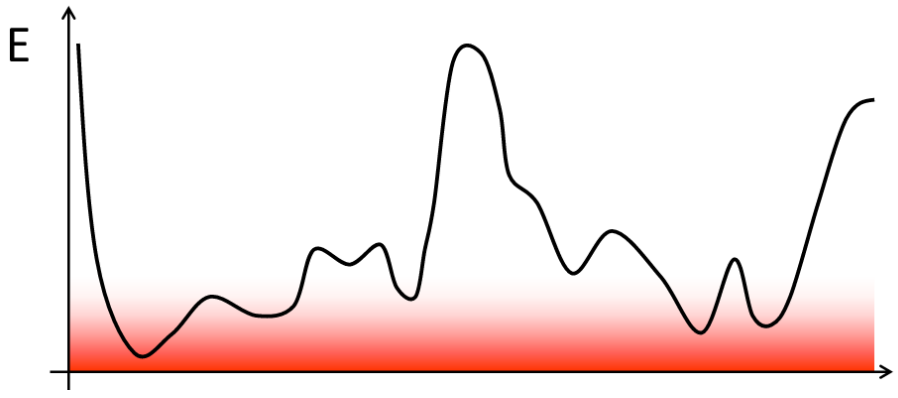
Similar to MCMC sample, but the goal is not representing the underlying mathematical function, but finding the global minimum using a metaheuristic approach.

Adds a new parameter, T.

$p(x_n, x_{n+1}) = min\left(exp(\frac{\Delta(f(x))}{T}), 1\right)$

This results in a rejection probability dependent on the temperature: If the temperature is very high the system will act as a random walker. If it's very low as a MCMC. #### Simulations in different coordinate space Cartesian: Is more flexible, but has many irrelevant conformations Torsion Space: Reduced to main variables, the search is efficient but not all conformations are possible

Influence of temperature in simulations



For simulations considering temperature, simulated annealing is used: First the simulation is started with a high temperature, where the system behaves as a random walker, as we lower temperature, the system starts to reject samples,

leading hopefully to the global minima, and being able to find the local minima feasible at different temperatures.

Overview