Transition Path Theory

Winterschool on
Theoretical Chemistry and Spectroscopy
9-13 December 2013, Han-sur-Lesse, Belgium

Bernd Ensing
Computation Chemistry
University of Amsterdam

Content

- Lecture 1: Markov processes
- Lecture 2: State clustering
- Lecture 3: Transition Path Theory
- Lecture 4: Sampling Reactive Events

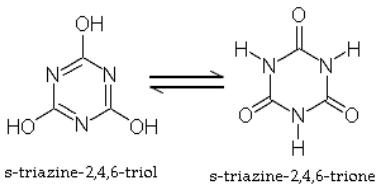
Friday, 13 December, 13

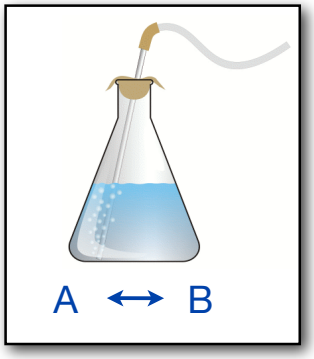
Stochastic processes

Chemical reaction:

$$A \leftrightarrow B$$

e.g. tautomerization reaction





At t=0

probability to measure molecule P(A) = 1.0
probability to measure molecule P(B) = 1.0-P(A) = 0.0

initial state distribution $S_0 = [1.0, 0.0]$

P(A) P(B)

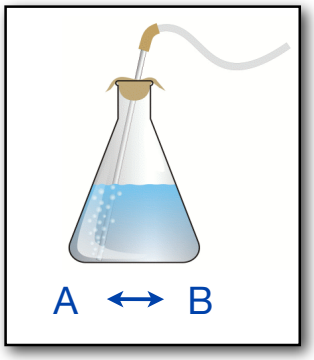
Stochastic processes

At t=0
initial state distribution $S_0 = [1.0, 0.0]$

e.g. a measurement every 30 minutes

Probability that A has converted to B after time τ

t(A→B) = 0.8
t(B→A) = 0.4



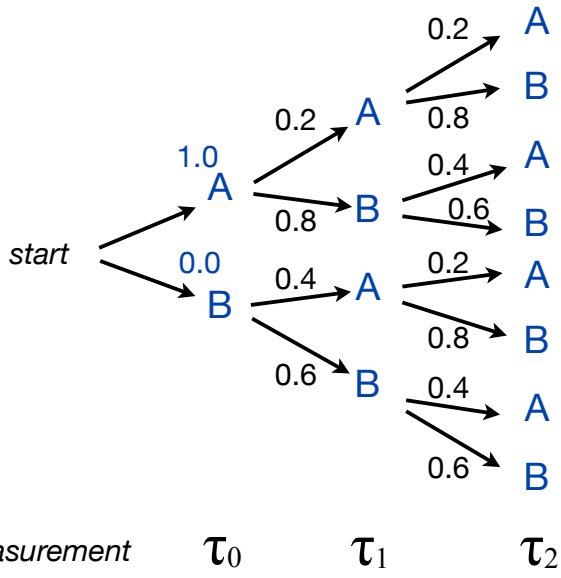
Transition matrix:

current state

$$T = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} \begin{matrix} A \\ B \end{matrix}$$

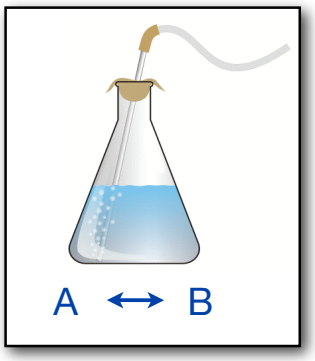
A B ← *next state*

Stochastic processes



$$S_0 = [1.0, 0.0]$$

$$\mathbf{T} = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}$$



$$S_1 = S_0 T = [1.0, 0.0] \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.2, 0.8]$$

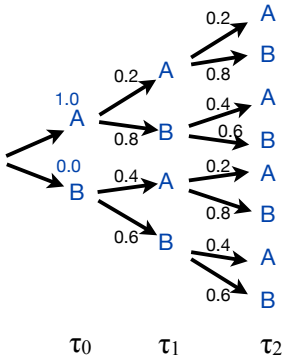
Stochastic processes

initial state distribution

$$S_0 = [1.0, 0.0]$$

transition matrix

$$\mathbf{T} = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}$$



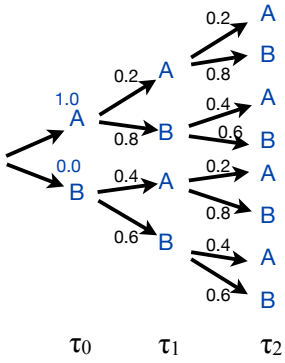
How does the distribution change in time?

$$S_1 = S_0 \mathbf{T} = [1.0, 0.0] \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.2, 0.8]$$

$$S_2 = S_1 \mathbf{T} = [0.2, 0.8] \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.36, 0.64]$$

$$S_3 = S_2 \mathbf{T} = [0.36, 0.64] \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.328, 0.672]$$

Stochastic processes



How does the distribution change in time?

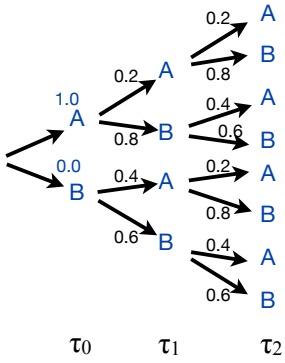
$$S_1 = S_0\mathbf{T} = [1.0, 0.0]\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.2, 0.8]$$

$$S_2 = S_1\mathbf{T} = [0.2, 0.8]\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.36, 0.64]$$

$$S_2 = S_0\mathbf{T}^2 = [1.0, 0.0]\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} =$$

$$= [1.0, 0.0]\begin{bmatrix} 0.36 & 0.64 \\ 0.32 & 0.68 \end{bmatrix} = [0.36, 0.64]$$

Stochastic processes



$$S_0 = [1.0, 0.0]$$

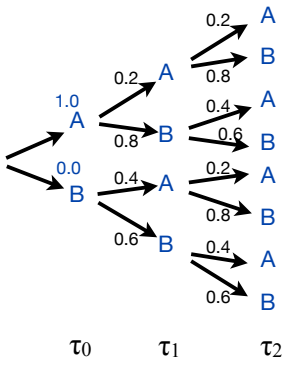
$$S_1 = S_0\mathbf{T} = [1.0, 0.0]\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.2, 0.8]$$

$$S_2 = S_1\mathbf{T} = [0.2, 0.8]\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.36, 0.64]$$

$$S_3 = S_2\mathbf{T} = [0.36, 0.64]\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [0.328, 0.672]$$

Does the distribution converge to a stable state?

Stochastic processes



$$S_0 = [1.0, 0.0]$$
$$S_1 = [0.2, 0.8]$$
$$S_2 = [0.36, 0.64]$$
$$S_3 = [0.328, 0.672]$$
$$S_4 = [0.3344, 0.6656]$$
$$S_5 = [0.33312, 0.66688]$$

$$S = \left[\frac{1}{3}, \frac{2}{3}\right]$$ *stationary matrix, steady state*

$$S\mathbf{T} = \left[\frac{1}{3}, \frac{2}{3}\right] \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = \left[\frac{1}{3}, \frac{2}{3}\right] = S$$

Markov process

Stochastic processes

- Poisson process:
  - number of events per time
  - no dependence between events
  - exponential decay
  - radioactive decay, earthquakes, jobs in printer queue
- Markov process:
  - stochastic process with simple relation between states X_1, X_2, X_3, X_4,...
  - probability to move to state X_{t+1} depends only on X_t (not on the past!)

formally: $P(X_{n+1=x}|X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = P(X_{n+1=x}|X_n = x_n)$

- probability that it rains in your garden (X discrete, time continue)
- brain activity (X continue, time continue)
- game of goose, DNA base at position *t* (X discrete, “time” discrete)
- daily solar activity (X continue, time discrete)

Andrej Markov

Andrej Andrejevitsj Markov (Russisch: Андрей Андреевич Марков, Rjazan, 14 juni 1856 - Petrograd, 20 juli 1922) was een Russisch wiskundige, naar wie de Markovketens, de Markovprocessen en de Markov-ongelijkheid zijn vernoemd. Zijn zoon (1903-1979) draagt dezelfde naam en is op het gebied van de algebra, topologie, mechanica en logica eveneens een bekend Russisch wiskundige.

Markov chain

A Markov chain is a discrete time Markov process

A regular Markov chain has a regular transition matrix

A transition matrix, **T**, is regular if some power of it, $\mathbf{T}^n$, has only positive (non-zero) entries

$$\mathbf{T} = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}$$ *regular*

$$\mathbf{T} = \begin{bmatrix} 0.0 & 1.0 \\ 1.0 & 0.0 \end{bmatrix}$$ *not regular*

$$\mathbf{T} = \begin{bmatrix} 0.3 & 0.7 \\ 1.0 & 0.0 \end{bmatrix}$$ *regular*

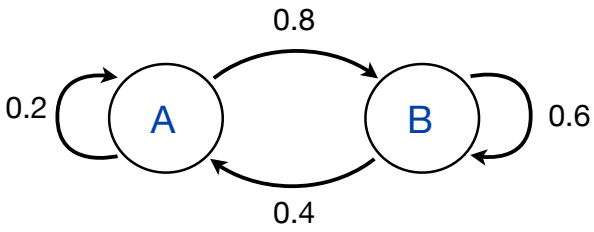
A regular Markov chain has a unique stationary state

Markov chain

Properties

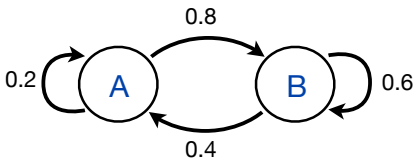
1. A regular Markov chain has a unique stationary state
2. The stationary state is found by solving: $S\,\mathbf{T} = S$
3. From any starting distribution S_0 the Markov chain will reach the stationary state
4. The matrices $\mathbf{T}^n$ approach a stationary matrix in which each row is the stationary state

Representation



Markov chain

Example, find the stationary state



$$S\,\mathbf{T} = S$$

$$S\,\mathbf{T} = [S_1, S_2] \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} = [S_1, S_2] = S$$

$$0.2S_1 + 0.4S_2 = S_1$$
$$0.8S_1 + 0.6S_2 = S_2$$

1) and 2) are linear dependent

$$S_1 + S_2 = 1$$

3) extra rule: probabilities add up to 1

$$S_1 = 1/3$$
$$S_2 = 2/3$$

Markov chain

Other properties:

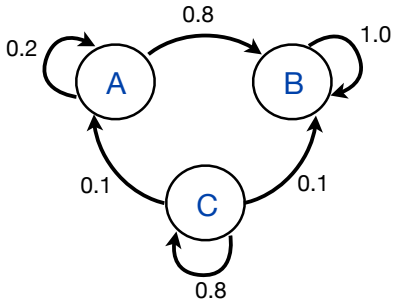
Time homogeneous Markov chain: transition matrix is time independent (otherwise time inhomogeneous).

Absorbing Markov chain: there is an absorbing state from which you cannot escape once you reach it.

n^{th} order Markov chain does not only depend on knowledge of the current state (=1st order) but also previous n-1 states.

A Markov chain is ergodic if there is a number of steps N in which each state can be reached from every other state (no loops, periodicity, absorbing state).

A Markov chain is reversible if it obeys detailed balance.



$$\mathbf{T} = \begin{bmatrix} 0.2 & 0.8 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

absorbing state: check diagonal for 1.0

Reversible Markov chains

Detailed balance

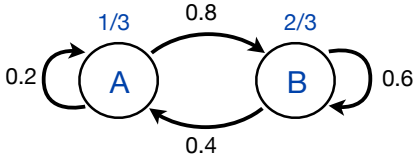
$$\rho_i \, P(X_{n+1} = j | X_n = i) = \rho_j \, P(X_{n+1} = i | X_n = j)$$

$$\rho_i \, t_{ij} = \rho_j \, t_{ji}$$

The probability be in state *i* times the probability to go to state *j* from state *i* is equal to be in state j times the probability to move from *j* to *i*.

The probability to be in state i, ρ_i , is an element of the equilibrium (stationary) state distribution S.

Example Reversible MC



$$\frac{1}{3} * 0.8 = \frac{2}{3} * 0.4$$

Detailed balance holds!

Reversible Markov chains

Example

$$\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

transition matrix

Detailed balance

$$\rho_i \, t_{ij} = \rho_j \, t_{ji}$$

$$S = [0.25, 0.25, 0.25, 0.25]$$

stationary state solution

Detailed balance does not hold!

See also a chain sample: ...ABCDABCDABCD...
Reverse B–>C transitions do not occur.

Mind you, if the stationary state solution is uniform so that

$$\rho_i = \rho_j \, \forall \, i \neq j \longrightarrow t_{ij} = t_{ji}$$

then the transition matrix should be symmetric for a reversible MC

For a reversible Markov chain it is not possible to determine the direction of the process from the observed state sequence alone.

Chapman-Kolmogorov

Transition matrix **T**
An element of **T** gives the transition probality:

$$t_{ij} = P(X_{n+1} = j | X_n = i)$$

The probability of going from state i to state j in *k* time steps is:

$$t_{ij}^{(k)} = P(X_{n+k} = j | X_n = i)$$

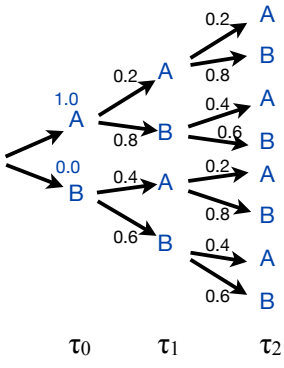
The Chapman–Kolmogorov equation says that for any *m* such that 0 < *m* < *k*,

$$p_{ij}^{(k)} = \sum_{r \in Z} p_{ir}^{(m)} p_{rj}^{(k-m)}$$

with the sum over all intermediate states *r* in the state space *Z*.

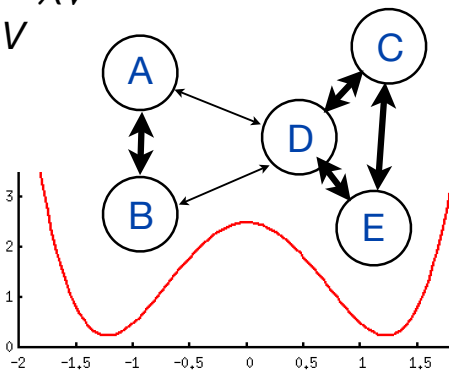
$$\mathbf{T} = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}$$

$$\mathbf{T}^{(k)} = \mathbf{T}^k$$



Spectral decomposition

Transition matrix **T** of dimension *d*
has *d* solutions of the linear equation: $\mathbf{T}\,V = \lambda V$
each with an eigenvalues λ with eigenvector *V*



Factorizing **T** in: $\mathbf{T} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$

in which **Λ** is the diagonal matrix of eigenvalues, which can be ordered:

$$1 = \lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_N| > 0$$

- Spectral decomposition allows to analyze how the system relaxes to equilibrium
- Eigenvalues close to 1 and significantly larger than the other eigenvalues indicate one or more slow relaxation times and metastability (rare events) between states.

Summary

part 1

A stochastic hopping between states is Markovian if the probability to reach the next state does not depend on the past (no memory)

- reversible Markov chains (detailed balance)
- ergodicity
- spectral decomposition (convergence to equilibrium)

Content

- Lecture 1: Markov processes
- Lecture 2: State clustering
- Lecture 3: Transition Path Theory
- Lecture 4: Sampling Reactive Events

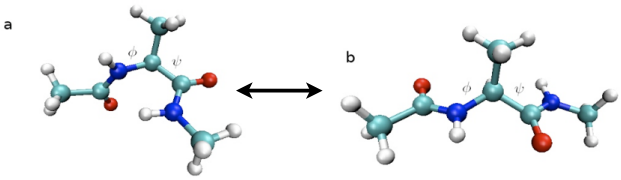
MSM of peptide dynamics

Markov State Modeling applied as analysis tool of the dynamics of a small peptide

- Analysis of molecular dynamics simulation
- Clustering based on structures
- Relation to free energy landscape
- Transition probability between clusters (states)

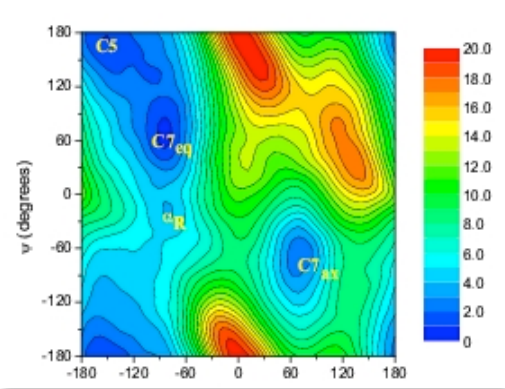
Is peptide dynamics Markovian?

How does a peptide (or protein) dynamically behave?



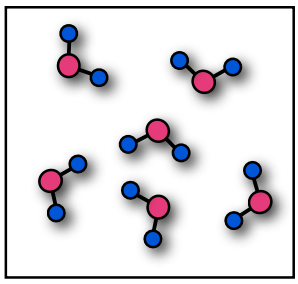
- can I distinguish characteristic structures/states?
- does it visit all states equally often?
- is there an order in visiting states?
- how do I know if it visited all possible states?
- can I compute the probability to visit a state?
- can I compute the probability to go from 1 state to any other state?

dipeptide



hexamer

FES of 12 torsions?



atom positions and velocities

$$r(t + \Delta t) = r(t) + v(t + \Delta t/2)\Delta t$$
$$v(t + \Delta t/2) = v(t - \Delta t/2) + \Delta t \frac{f(t)}{m}$$

Molecular Dynamics

Newtonian dynamics

$$f = m \cdot a$$
$$f = -\nabla V(r)$$

$$\Delta t \approx 1\,\text{fs} = 0.000000000000001\,\text{sec}$$

forcefield

$$V(\mathbf{r}) = \sum_{bonds} k_r(r - r_{eq})^2 + \sum_{angles} k_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{1}{2}\nu_n(1 + cos(n\phi - \phi_0)) + \sum_{i<j} \left(\frac{a_{ij}}{r_{ij}^{12}} - \frac{b_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\epsilon r_{ij}} \right)$$

bonds bends torsions non-bonded

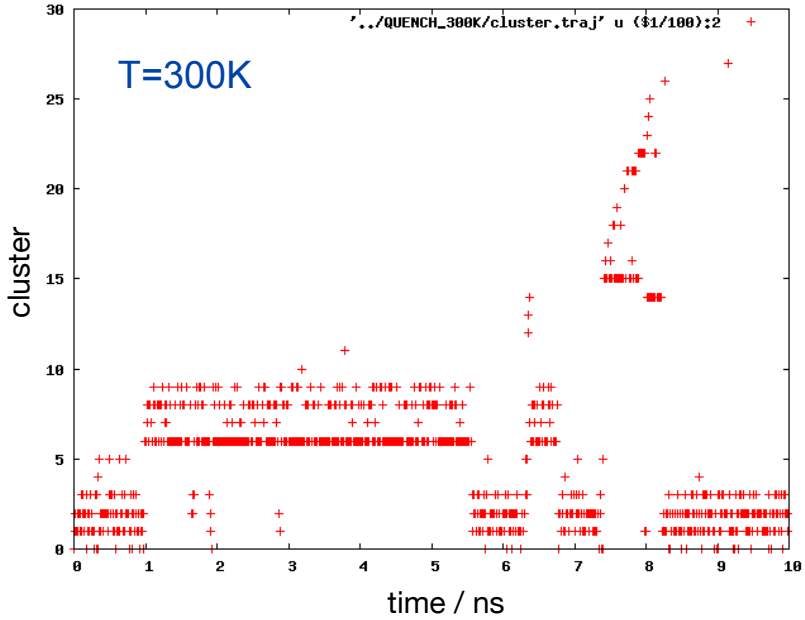
Cluster analysis

- 3x 10-20 ns simulations at T= 300, 400, and 500 K
- take a frame every 10 ps and quench in to zero Kelvin
- cluster quenched frames based on torsion angles

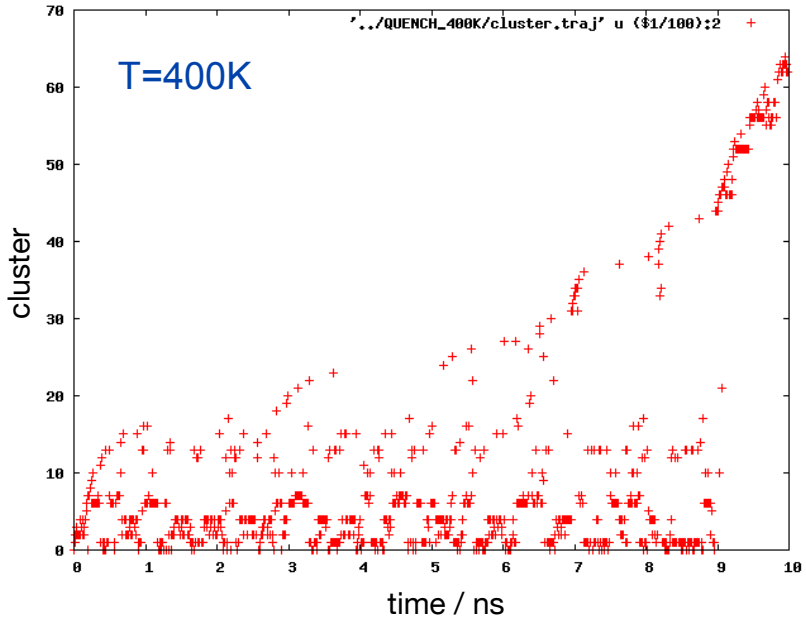
Clustering of structures based on resemblance between observed geometries in the trajectory,
e.g. RMSD(R_i,R_j) < TOLL

many flavors of clustering methods exist
typically involves an iterative procedure over the trajectory

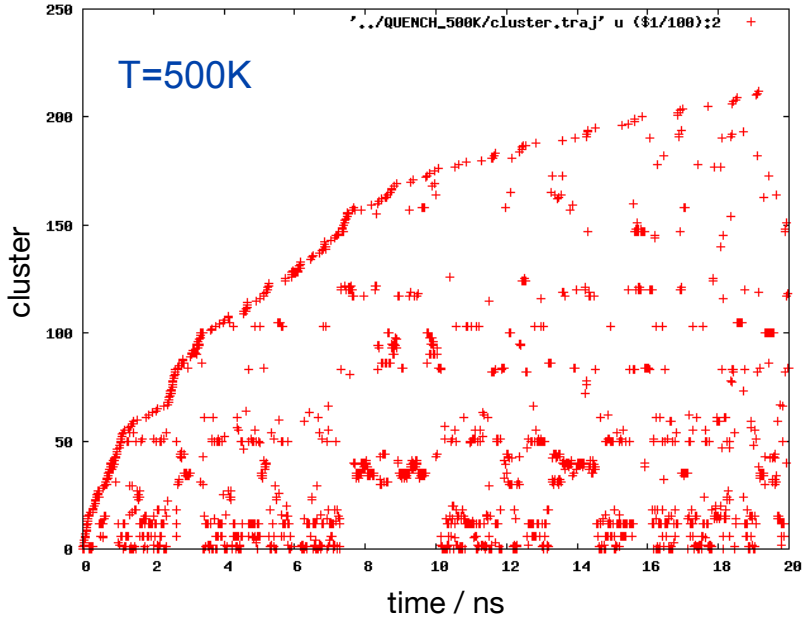
Cluster analysis



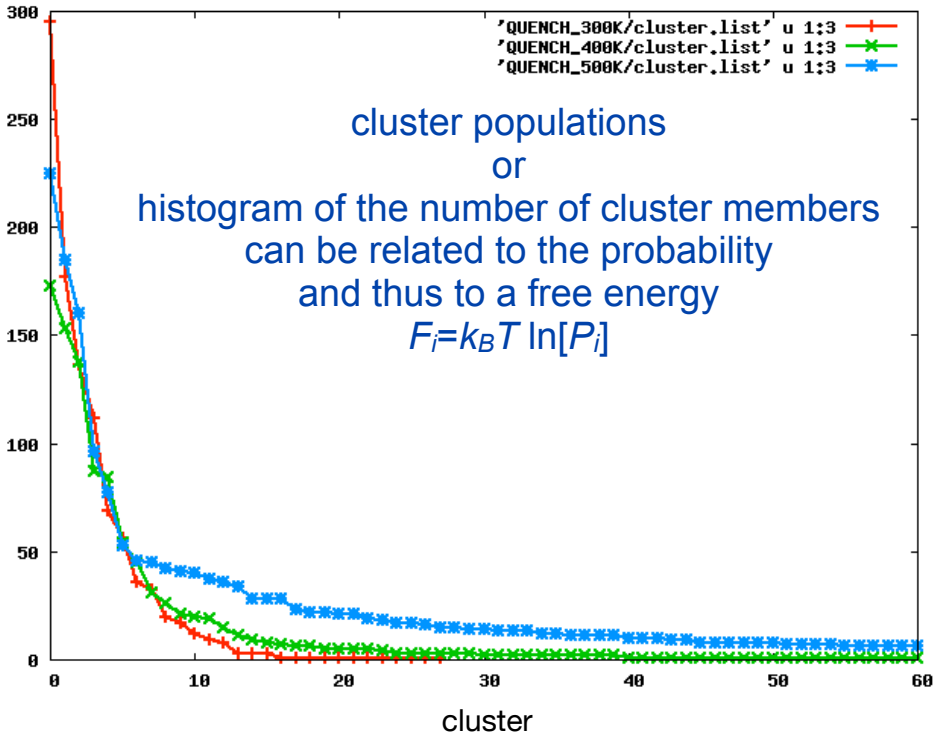
Cluster analysis



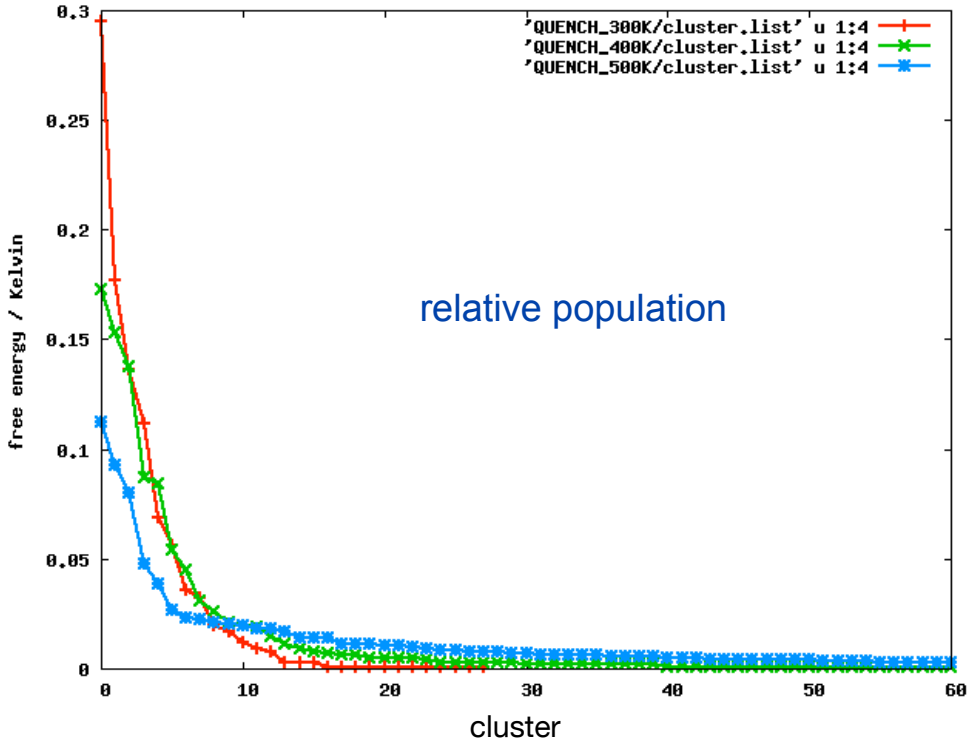
Cluster analysis



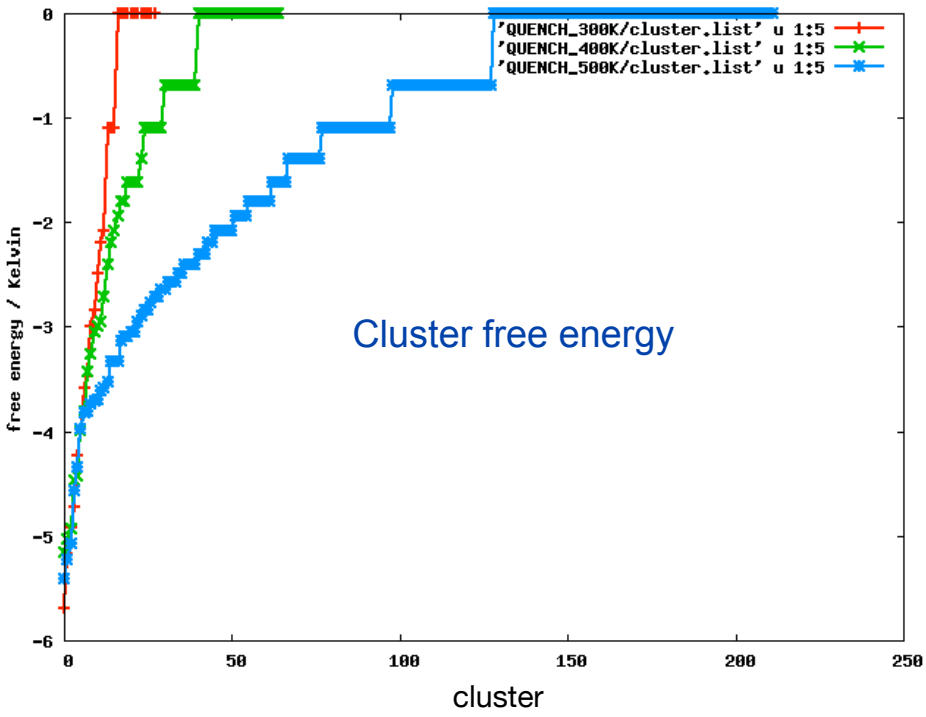
Cluster analysis



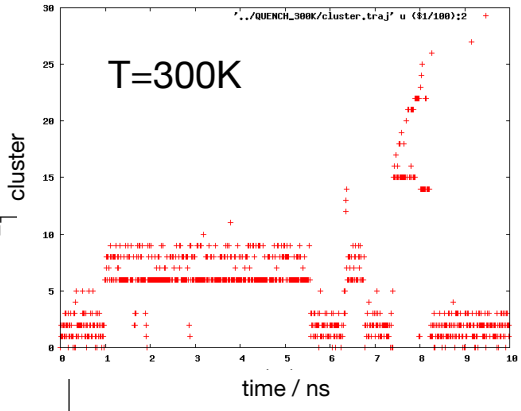
Cluster analysis

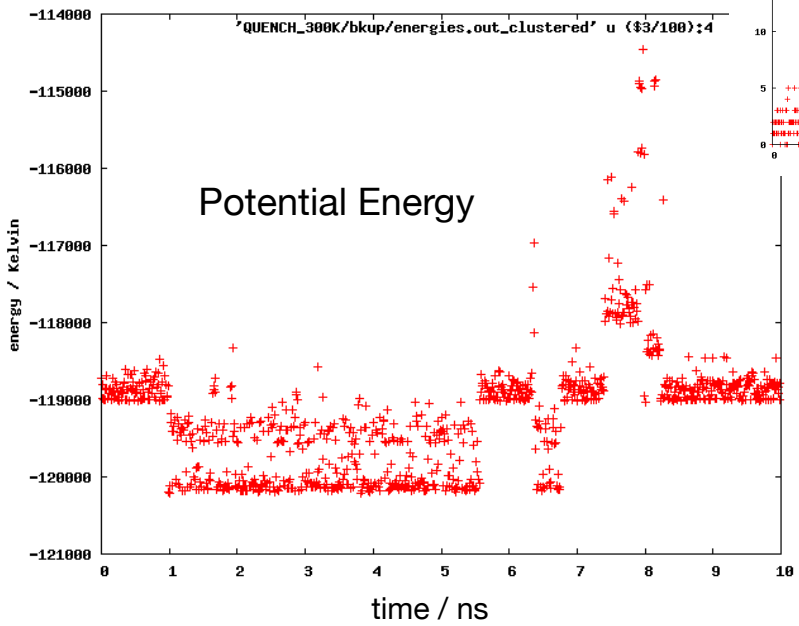


Cluster analysis

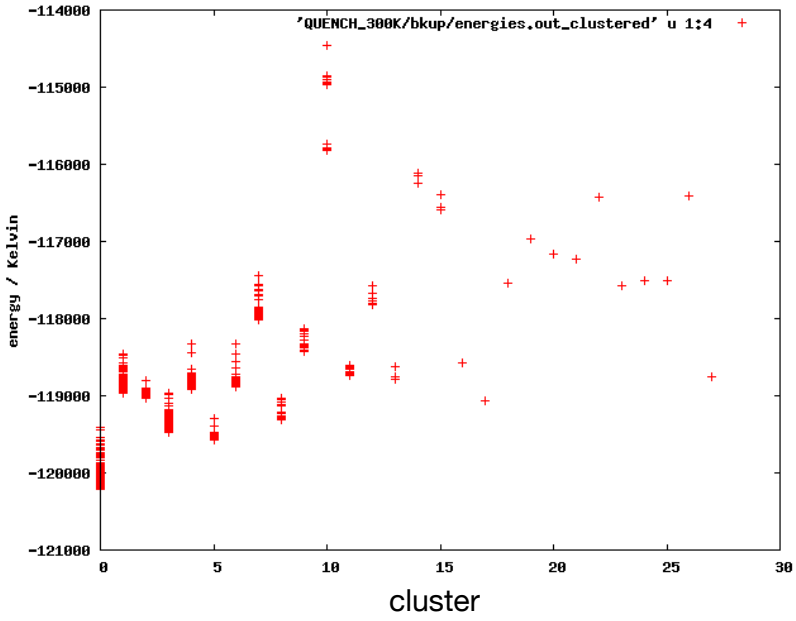


Do clusters with most members have the lowest potential energy?



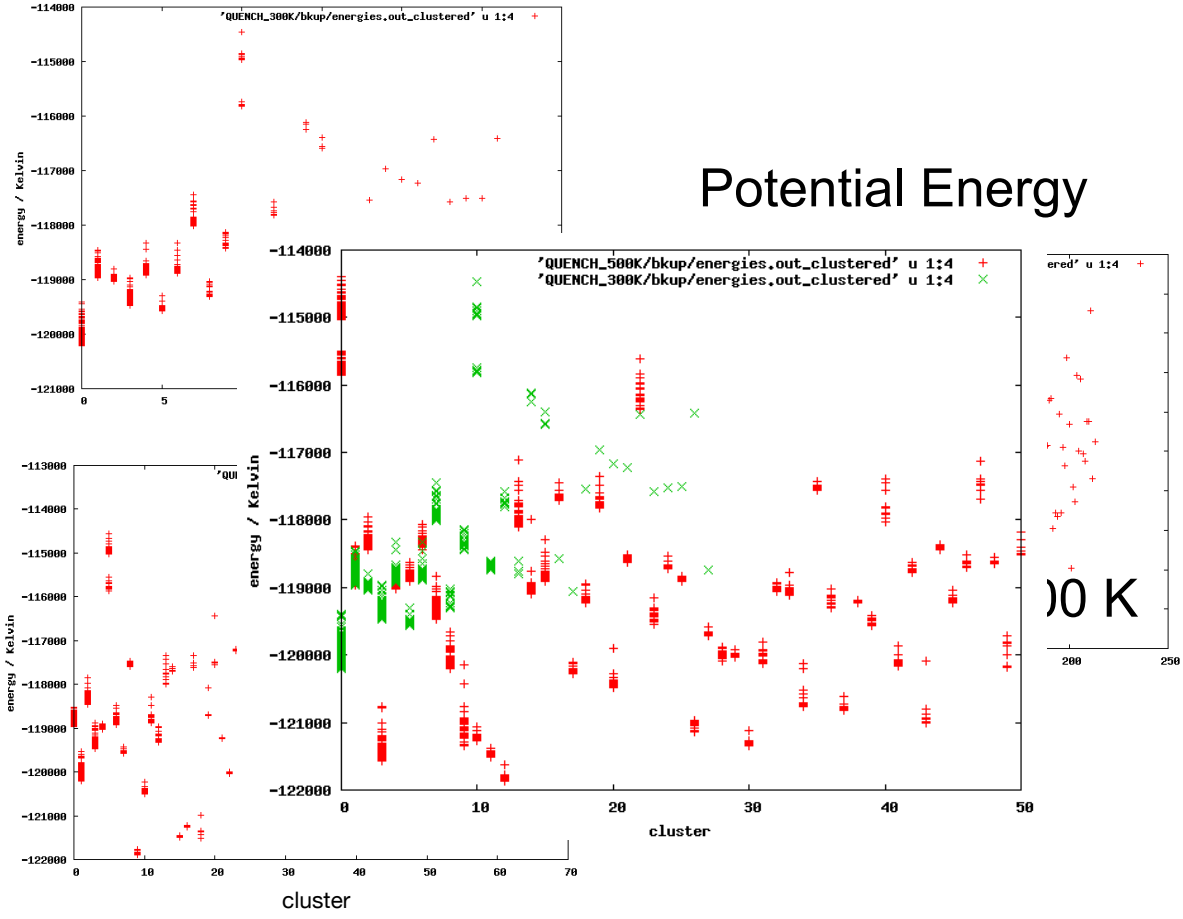


T=300K Potential Energy

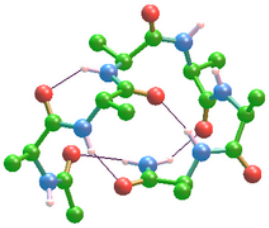


There is a correlation (cluster 1 has lowest E, but cluster 2 not second lowest).
Correlation not perfect, total energy not full story...

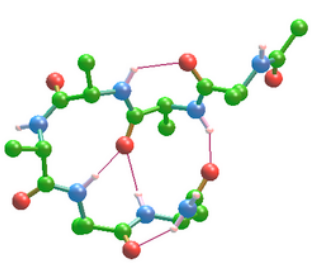
Potential Energy



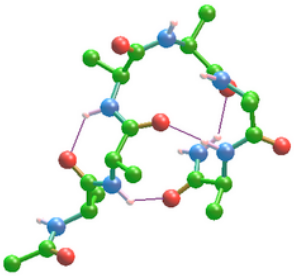
Cluster structures



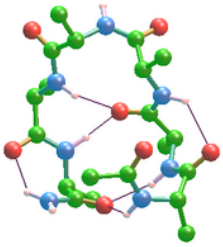
cluster 0



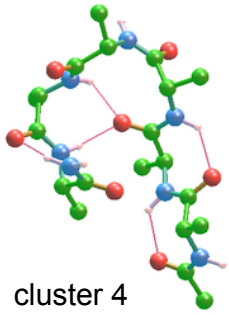
cluster 1



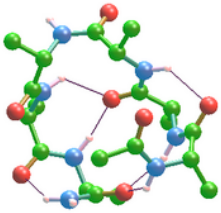
cluster 2



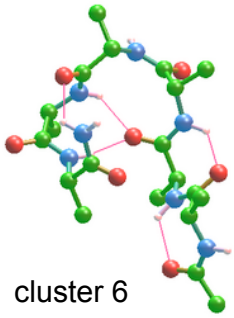
cluster 3



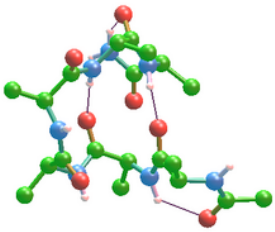
cluster 4



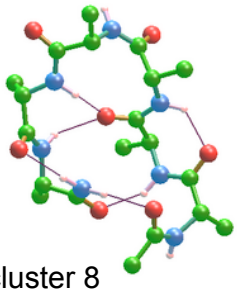
cluster 5



cluster 6



cluster 7



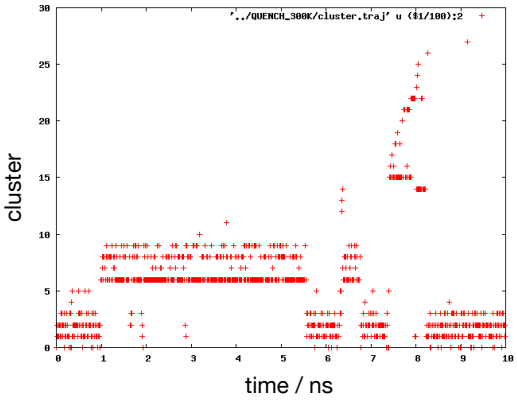
cluster 8

Cluster Dynamics

Transition matrix

Average Connection Frequency:

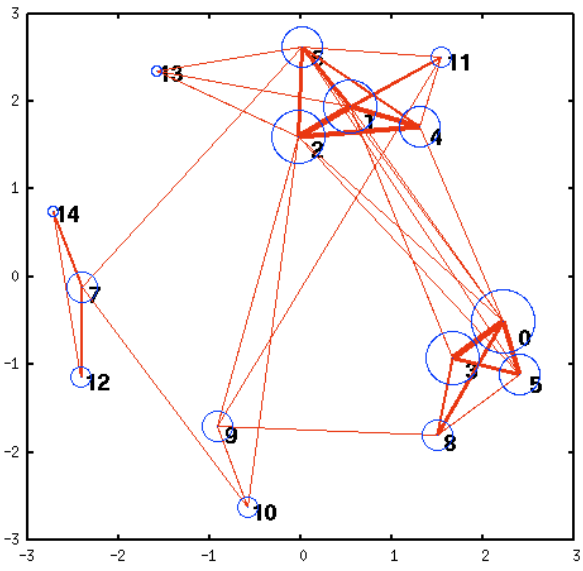
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	1	1	1	127	2	78	1	0	22	0	0	0	0	0	0
1		0	127	2	70	0	25	0	0	0	0	6	0	2	0
2			0	0	37	1	18	0	0	2	1	5	0	2	0
3				0	0	17	0	0	12	0	0	0	0	0	0
4					0	0	12	0	0	0	0	3	0	0	0
5						0	1	0	3	0	0	0	0	0	0
6							0	1	0	0	0	3	0	2	0
7								5	0	0	1	0	7	0	5
8									0	1	0	0	0	0	0
9										2	2	1	0	0	0
10											0	0	0	0	0
11												0	0	0	0
12													0	0	1
13														0	0
14															0



Free Energy Barriers:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	-0.00	-0.00	-0.00	-4.84	-0.69	-4.36	-0.00	*	-3.09	*	*	*	*	*	*
1		*	-4.84	-0.69	-4.25	*	-3.22	*	*	*	*	-1.79	*	-0.69	*
2			*	*	-3.61	-0.00	-2.89	*	*	-0.69	-0.00	-1.61	*	-0.69	*
3				*	*	-2.83	*	*	-2.48	*	*	*	*	*	*
4					*	*	-2.48	*	*	*	*	-1.10	*	*	*
5						*	-0.00	*	-1.10	*	*	*	*	*	*
6							*	-0.00	*	*	*	-1.10	*	-0.69	*
7								-1.61	*	*	-0.00	*	-1.95	*	-1.61
8									*	-0.00	*	*	*	*	*
9										-0.69	-0.69	-0.00	*	*	*
10											*	*	*	*	*
11												*	*	*	*
12													*	*	-0.00
13														*	*
14															*

Network diagram



circle size: free energy of cluster

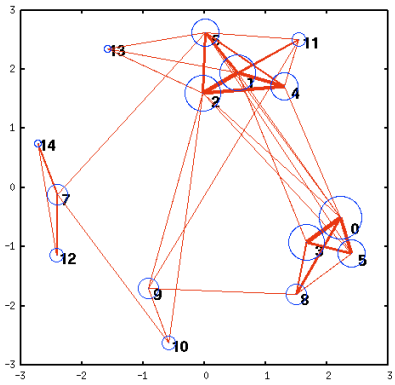
line thickness: transition probability

We can think of the peptide dynamics as a Markov chain at different levels:

- 14 small states
- 4 combined states

A Markov model

Are we allowed to apply a Markov model for analysis?
(mind you, the molecular dynamics was deterministic (not stochastic)).



We assume that the system jumps from state to state in a stochastic manner

- i.e. there is a separation of timescales such that the system stays long enough in metastable state to loose memory of the previous state ($\tau_{ij} >> \tau_{relax}$)

If the cluster analysis is not converged, we may find that the Markov chain is not reversible...

The Markov model can be stochastically sampled using Langevin dynamics, kinetic Monte Carlo, Fokker-Planck,....

Summary

part 1

A stochastic hopping between states is Markovian if the probability to reach the next state does not depend on the past (no memory)

- reversible Markov chains (detailed balance)
- ergodicity
- spectral decomposition (convergence to equilibrium)

part 2

Markov State Modeling applied as analysis tool of the dynamics of a small peptide

- Analysis of molecular dynamics simulation
- Clustering based on structures
- Relation to free energy landscape based on cluster size
- Transition probability between clusters (states)
- Using the MSM the dynamics can be sampled in a coarse-grain manner, e.g. using kinetic Monte-Carlo