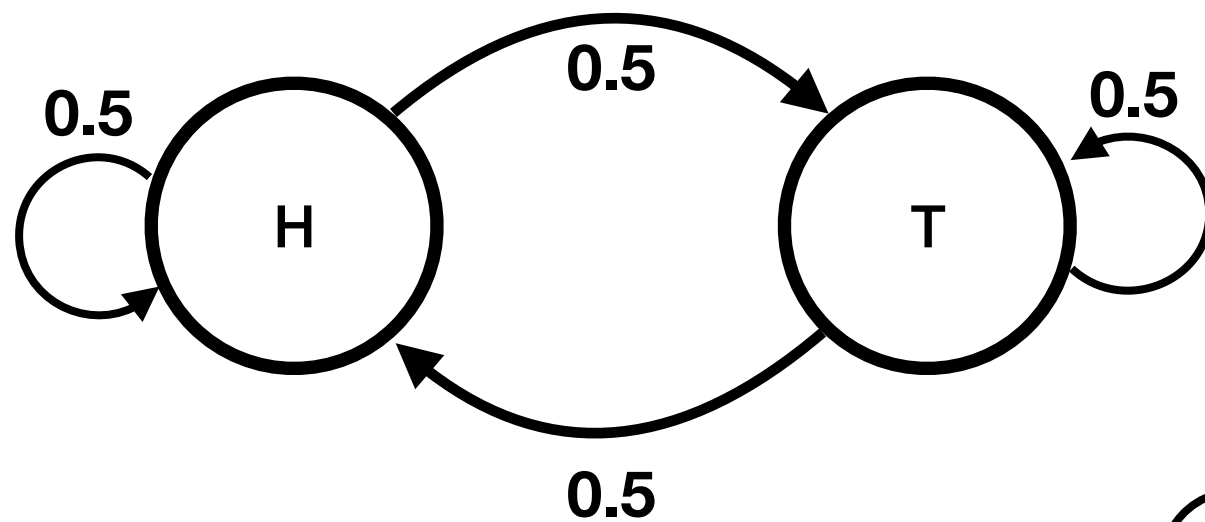


Introduction to Markov State Models

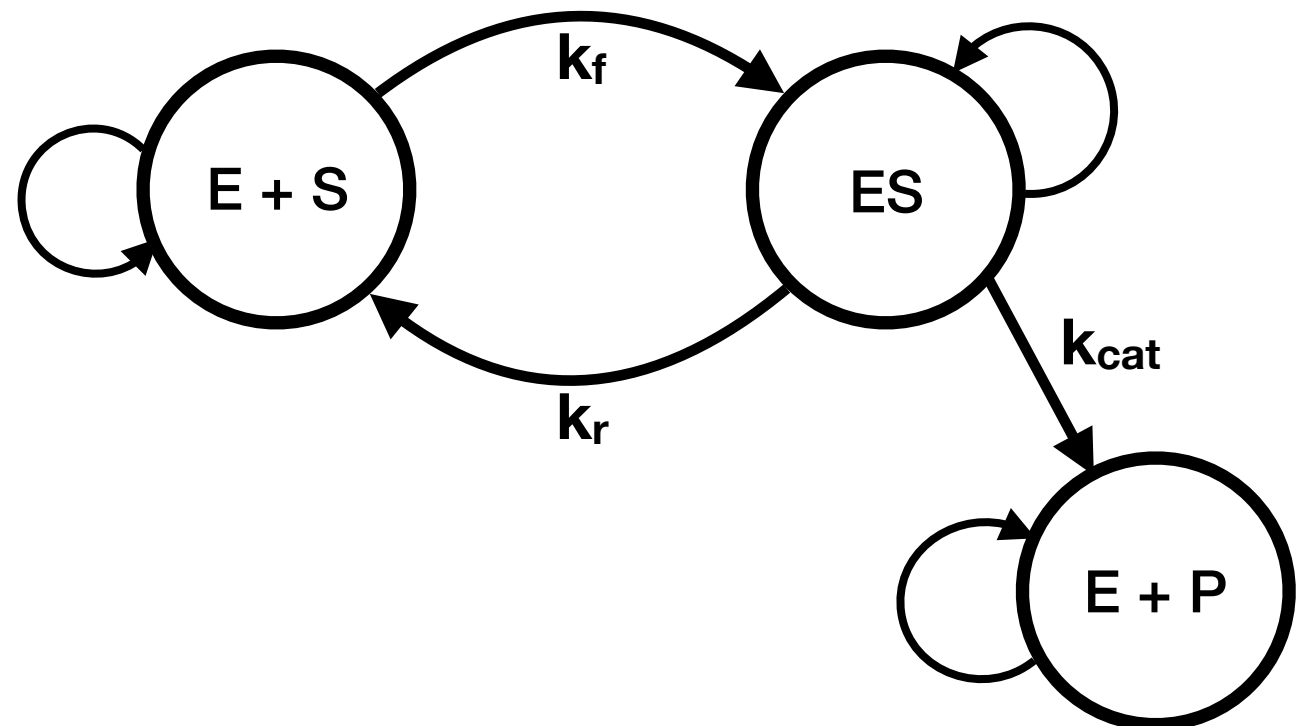
Drug Computing — Mobley Lab — UCI

Markov processes describe the dynamics of systems that have no “long term memory”

Flipping a coin

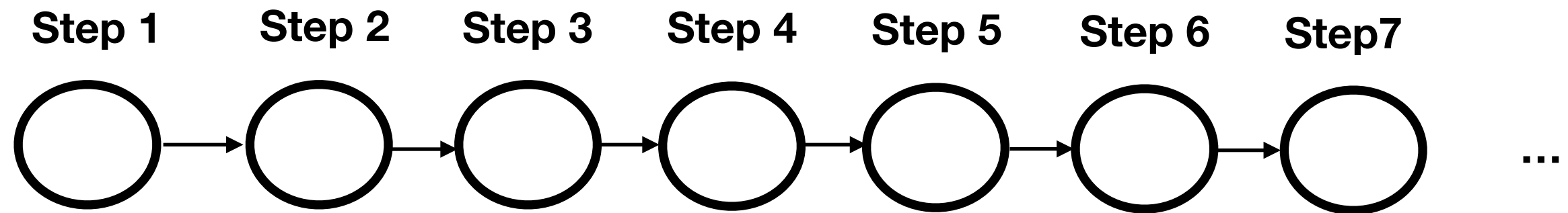


Michaelis Mentin Kinetics

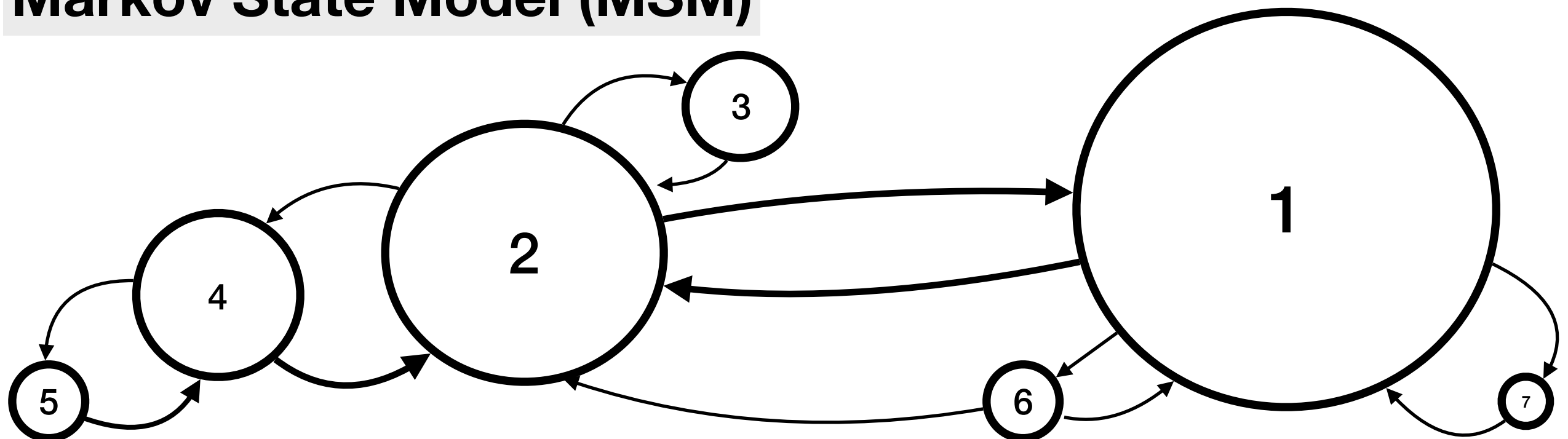


Markov State Models (MSMs) reveal kinetics of configurational transitions

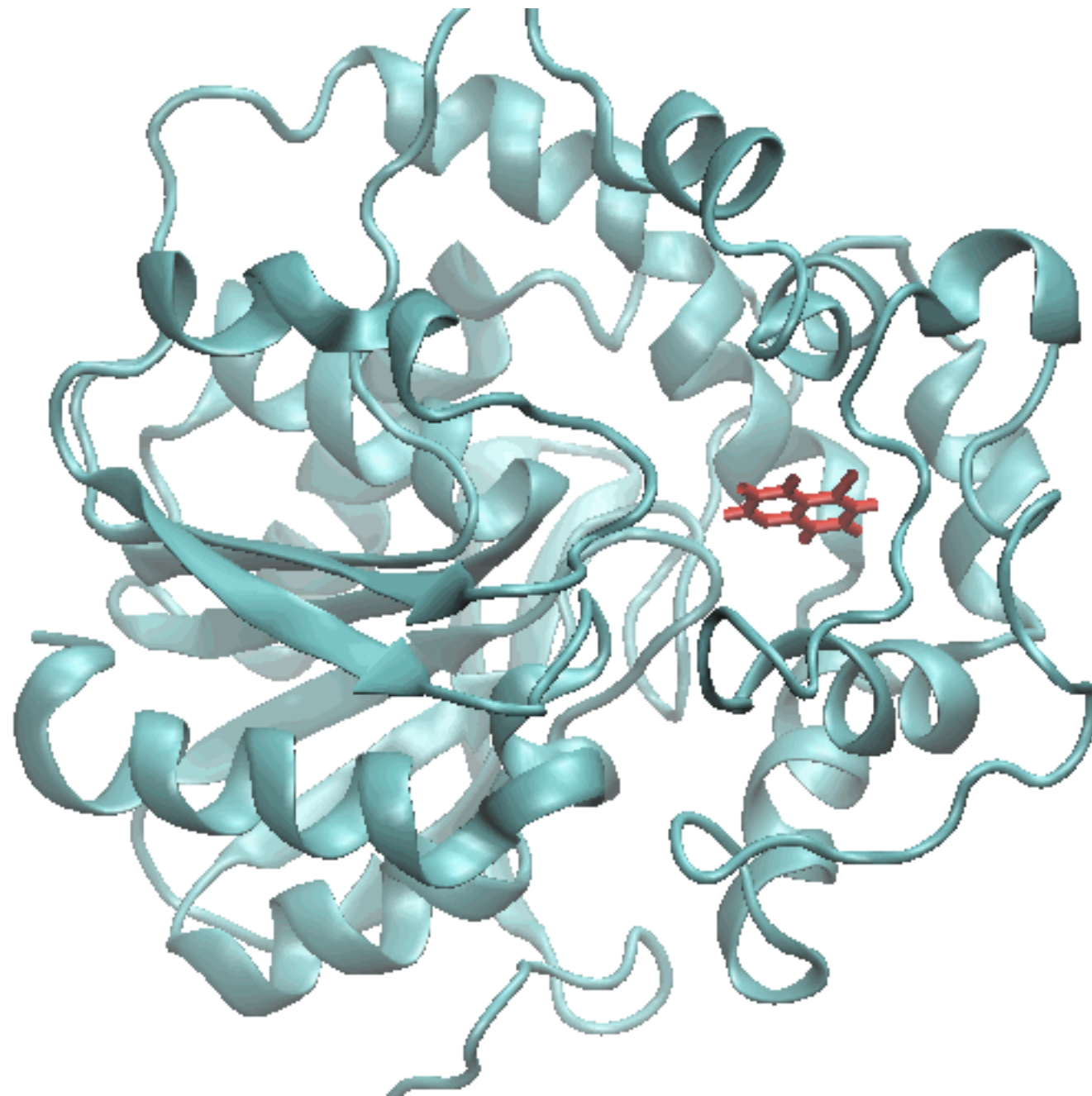
Highly Disordered Trajectory



Markov State Model (MSM)



Markov State Models (MSMs) reveal kinetics of configurational transitions



Building a Markov State Model

(1) Identify Configuration Microstates

- Dimensionality Reduction**
- Free Energy**
- PCA, TICA, etc.**

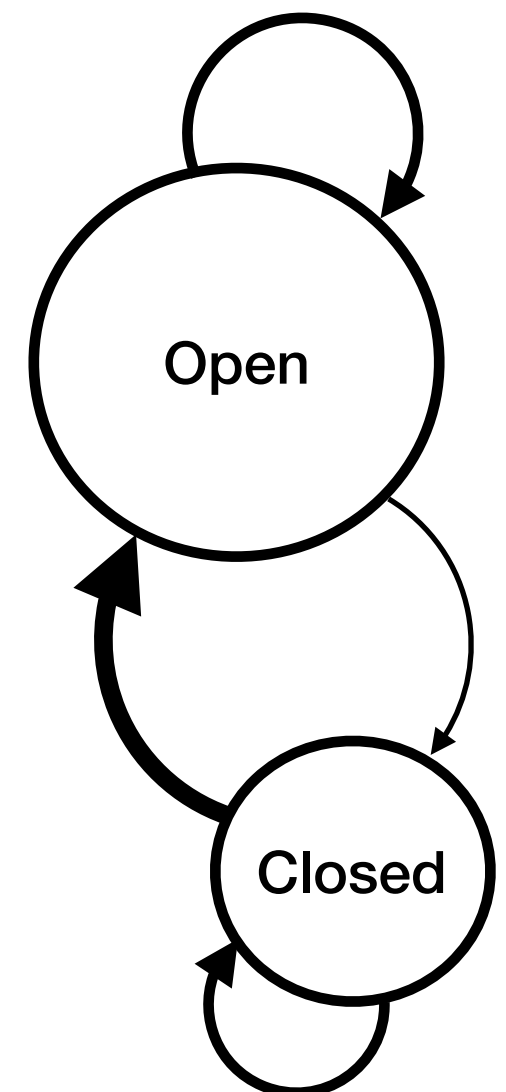
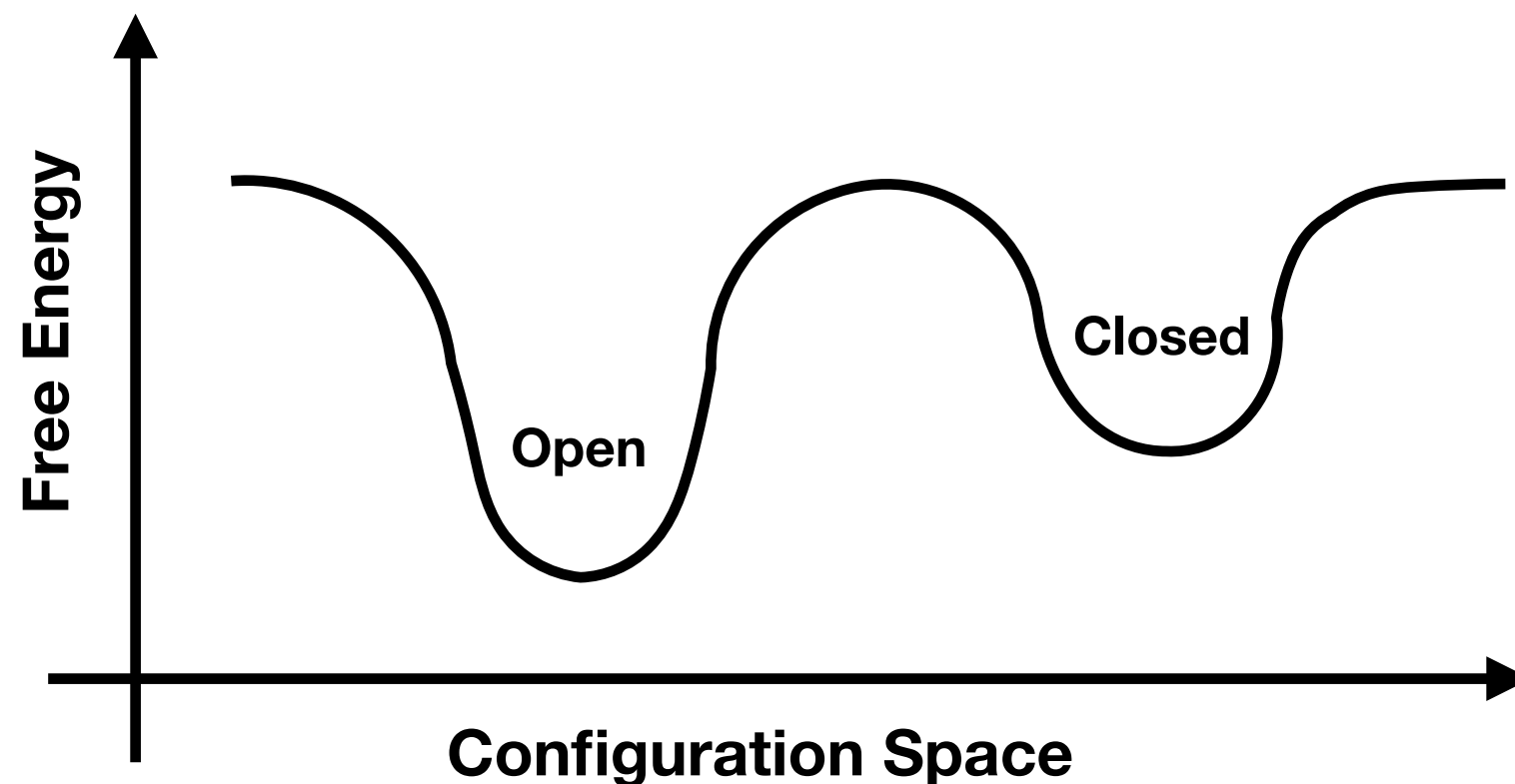
(2) Construct Transition Matrix Between States

(3) Test for Self-Consistency

Building a Markov State Model

(1) Identify Configuration Microstates from basins in the Free Energy Landscape

- Dimensionality Reduction



Dimensionality Reduction: Principle Component Analysis (PCA) and Time-independent Component Analysis (TICA)

**Huge Amount of Data ==> A Few Dimensions of
Variance**

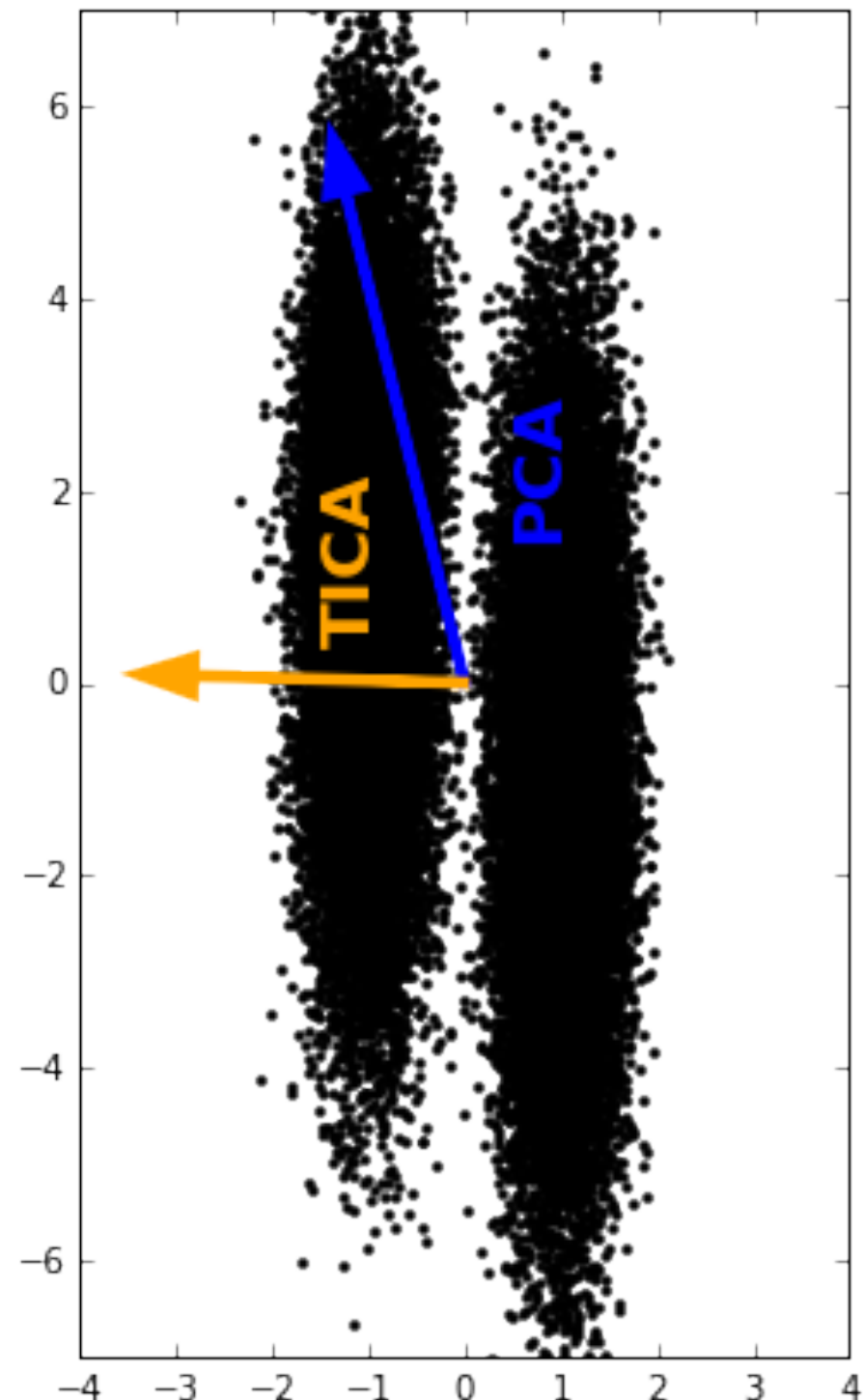
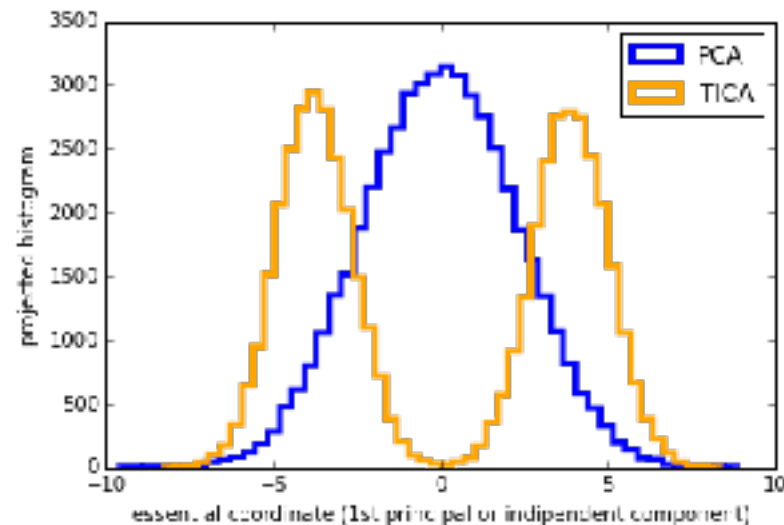
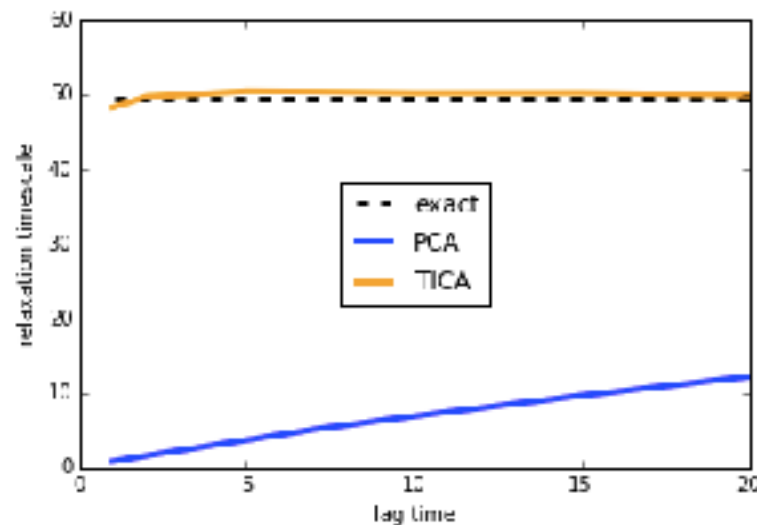
Brute Force Method From Trajectory:

- Transition Matrix between states defined by positions of every atom**
- N atoms, T time steps —> 3NT columns**
- Ex:**
 - 2500 atoms, 50,000,000 time steps
= 375 billion elements**

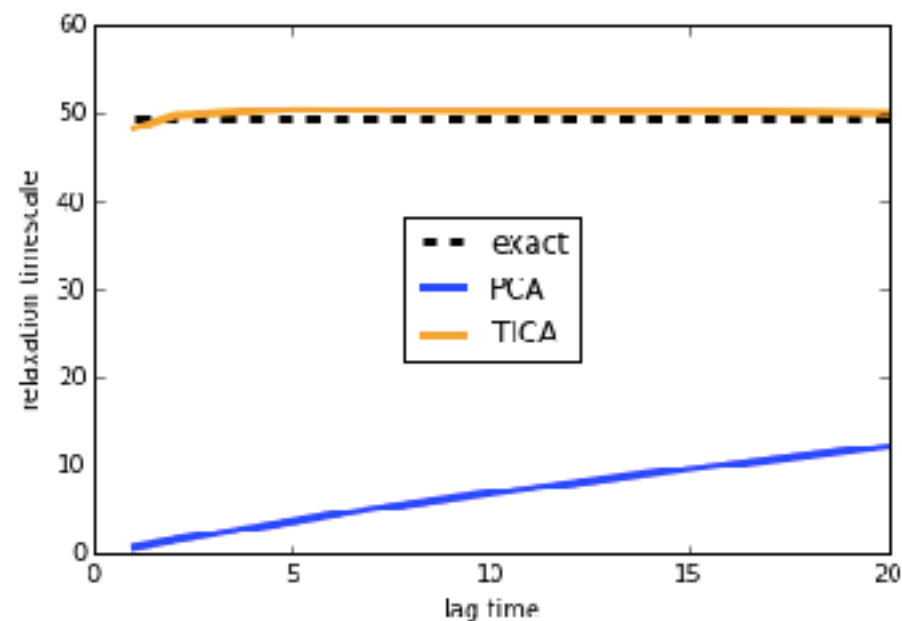
Dimensionality Reduction: Principle Component Analysis (PCA) and Time-independent Component Analysis (TICA)

PCA finds the direction of maximum variance in the data

TICA identifies the slowest reaction coordinates (maximum autocorrelation)



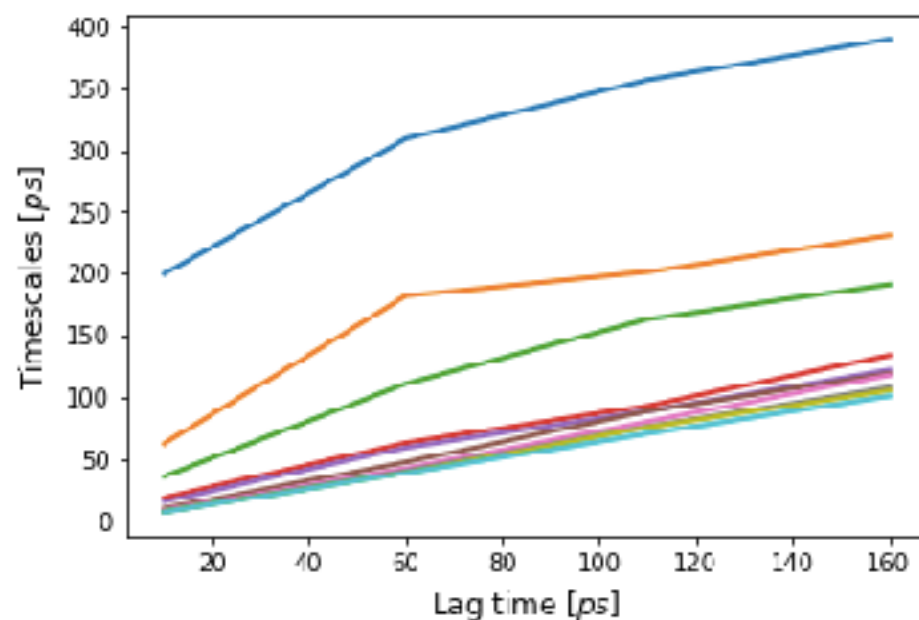
Testing the dimensionality reduction: Implied Timescales



Perfectly Markovian Process

$$t_i(k\tau_0) = t_i = -\frac{\tau_0}{\ln \lambda_i(\tau_0)}$$

http://docs.markovmodel.org/lecture_implied_timescales.html



Imperfect Markov Model

$$t_i(k\tau_0) = -\frac{k\tau_0}{\ln \lambda_i(k\tau_0)}$$

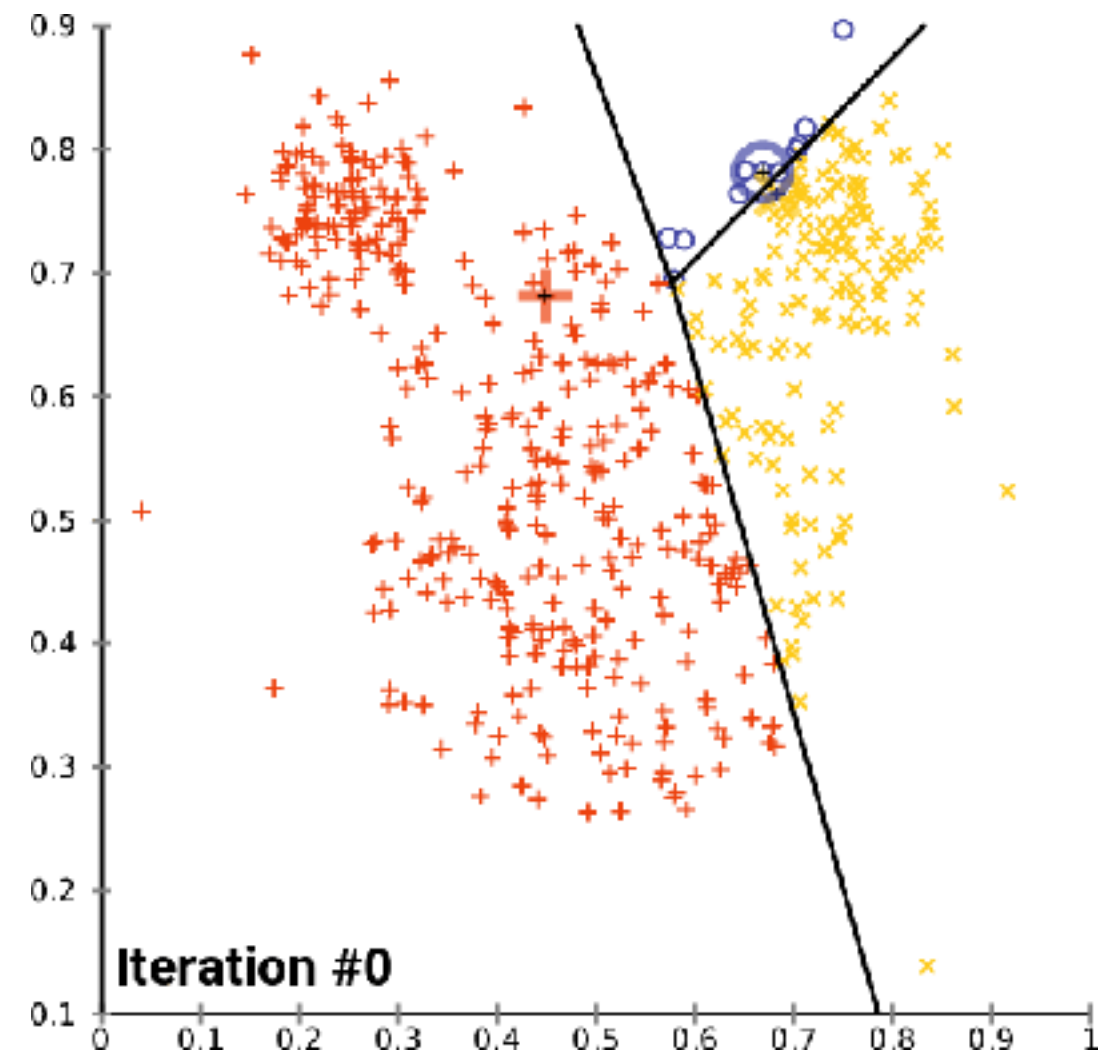
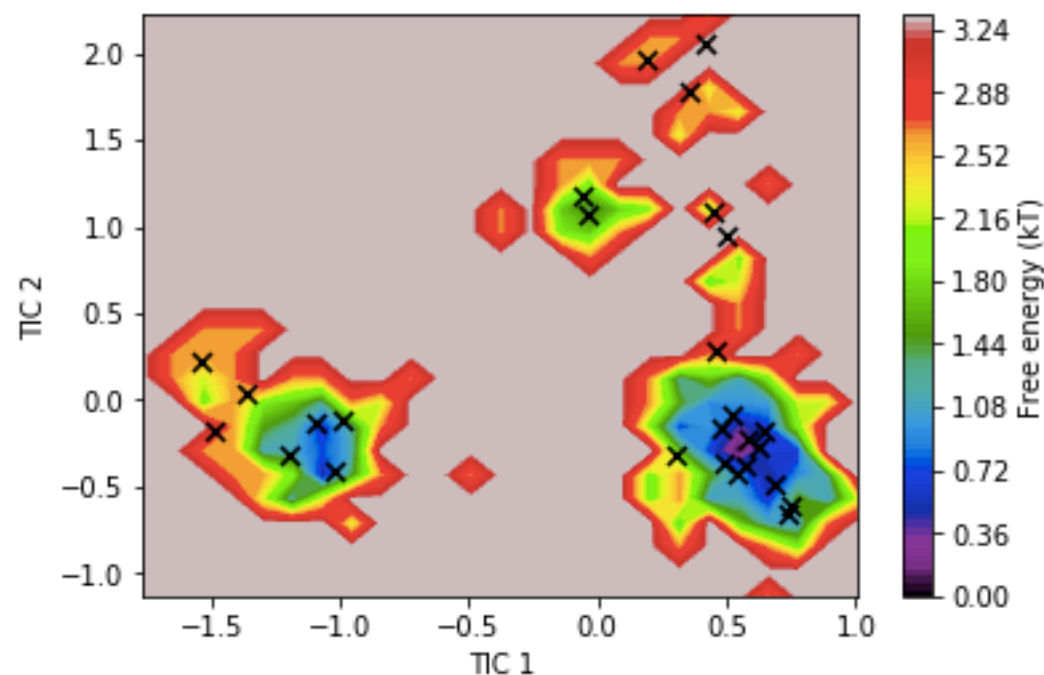
Choose a time scale at which the model appears to start behaving Markovian

Using the dimension reduction to identify macrostates of the system

PCCA and K-means Clustering

K-means Clustering:

- Identifies clusters based on distance in TICA space

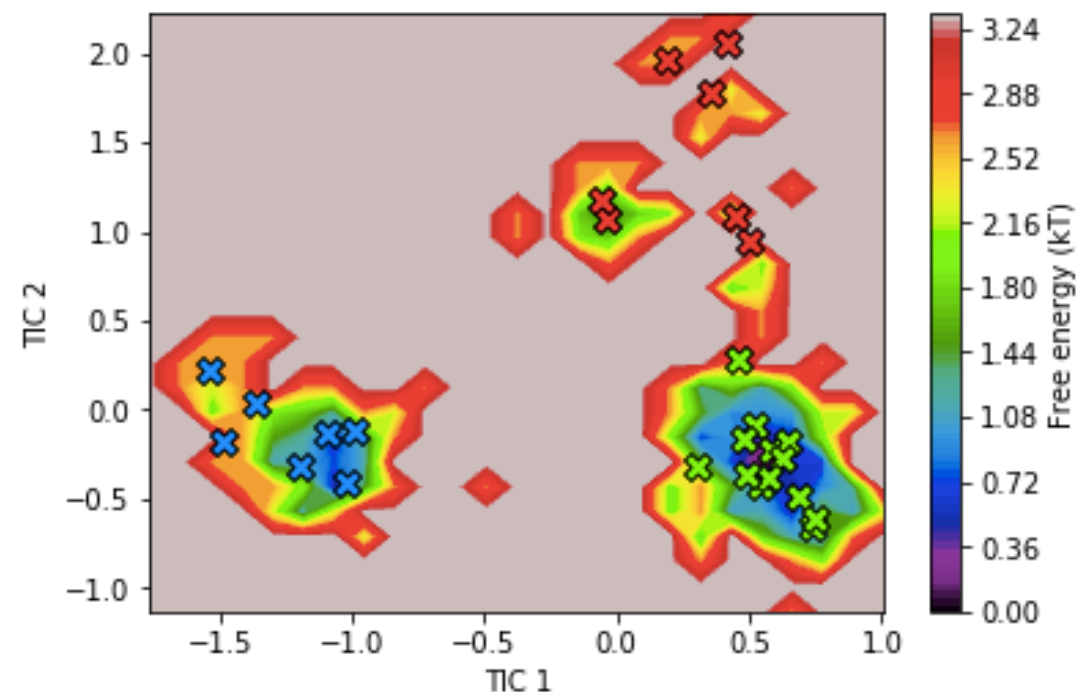
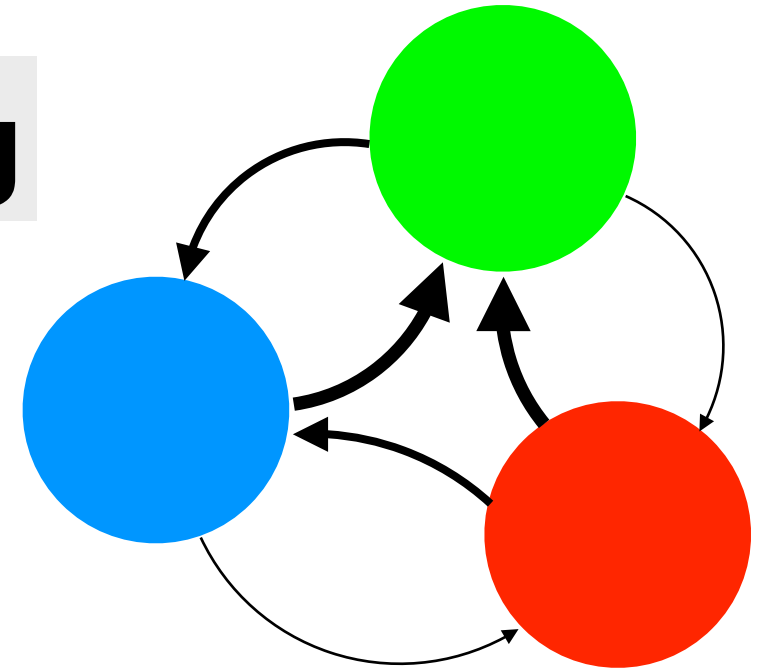


Using the dimension reduction to identify macrostates of the system

PCCA and K-means Clustering

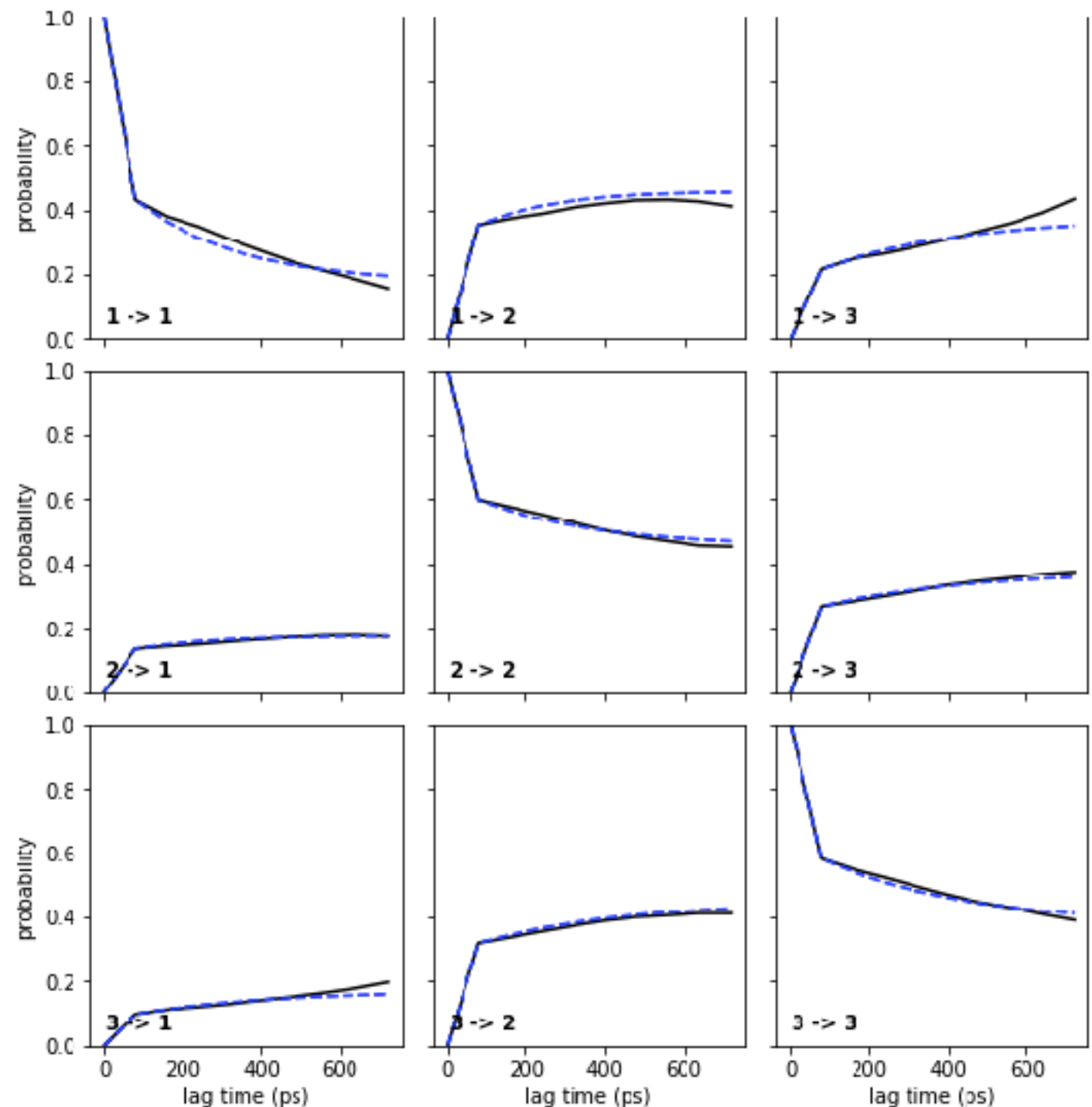
PCCA: Perron-Cluster Cluster Analysis

- Uses the eigenvector-eigenvalue decomposition to identify meta-stable states
- Splits the available conformational state space in order of longest relaxation time



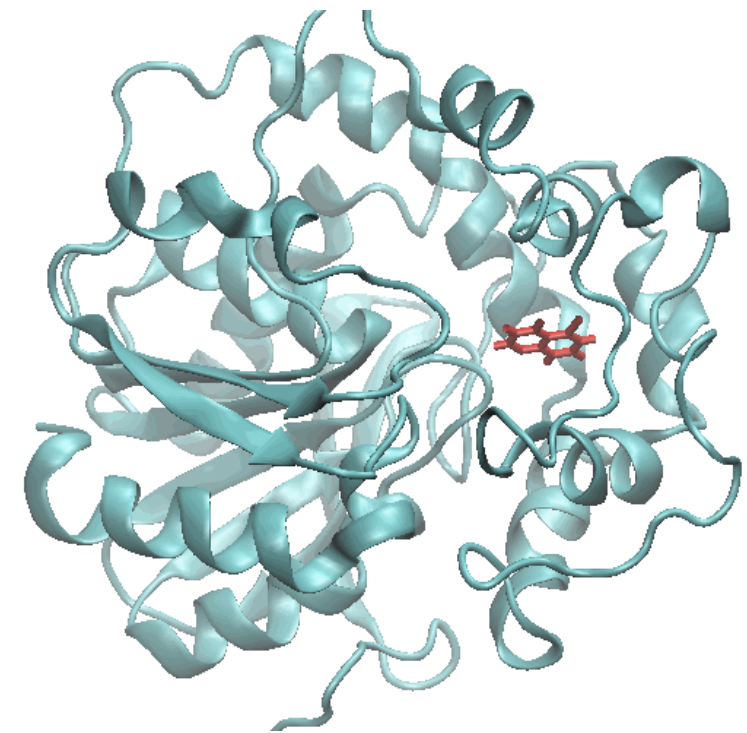
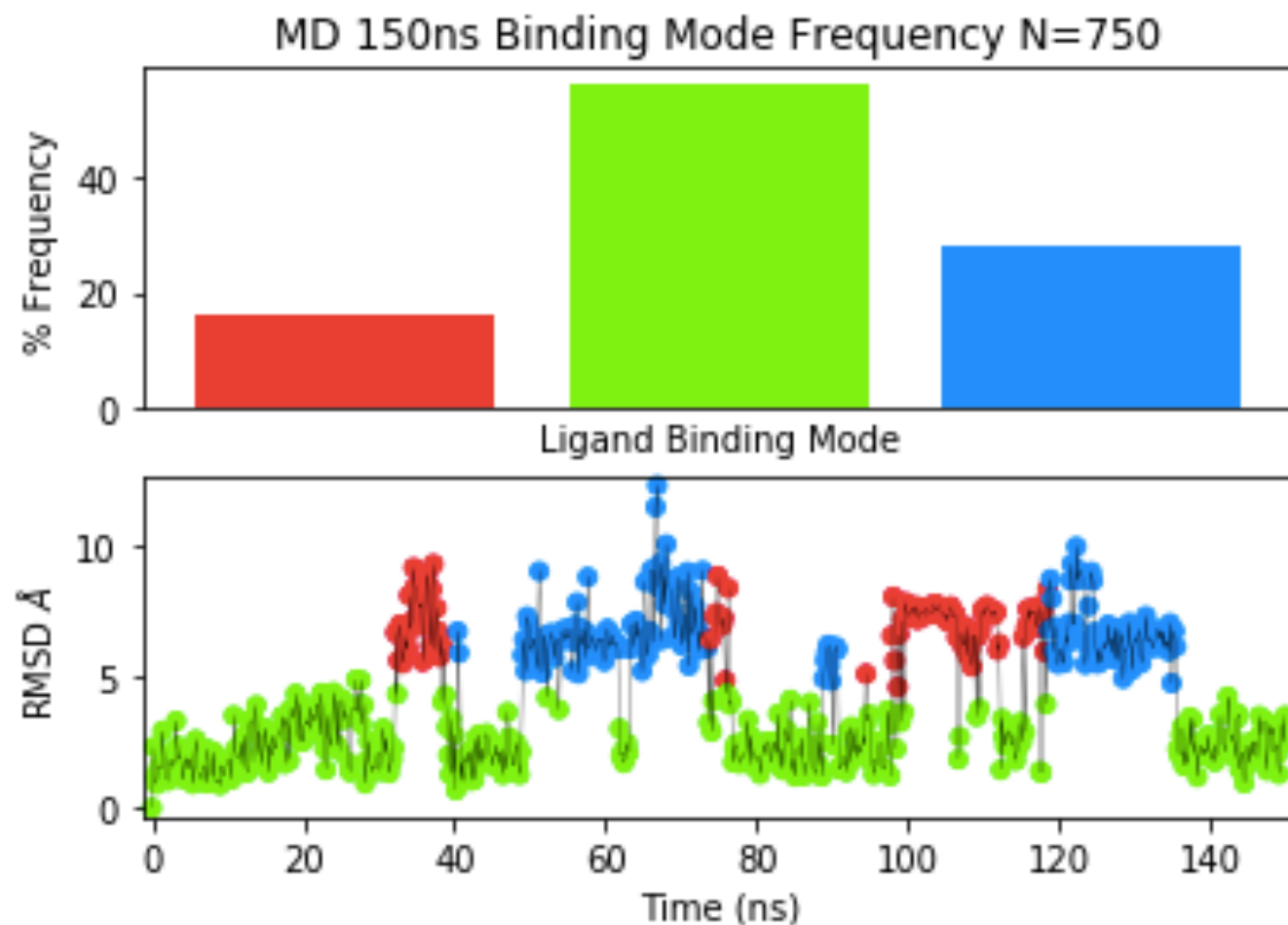
Testing the Markov Model: Chapman-Kolmogorov test

**How well do the
interconversions
between the
course-grained
macro-states
of the model
compare to the
real trajectory?**



Once we have a MSM, what can we do with it?

Identify Dominant Binding Modes of Ligands



Once we have a MSM, what can we do with it?

Identify Dominant Protein Folding Pathways

