Analysing NaCL umbrella sampling data using TRAM

February 24, 2022

1 Dataset

Given is a dataset of a Na-Cl-dimer in water; the dataset includes timeseries of the distance between the Na and Cl atom, which spans from a bound (smallest distances) to an unbound (largest distances) state.

The dataset contains data from 60 biased simulations (governed by a harmonic umbrella potential), which incrementally pull the ions apart (by spacing bias potentials along along the distance axis), and 20 unbiased simulations started at a Na-Cl-distance of approximately 3.5 Angstrom. The umbrella sampling parameters and the k_BT value are also given.

All distances in the dataset are in Angstrom, energies in kcal/mol, and the trajectory timestep is 1 ps.

Download the dataset from the ftp using the following code snippet:

```
with np.load(mdshare.fetch('pyemma-tutorial-us-nacl.npz', working_directory='data')
    ) as fh:
    us_trajs = [fh['us_traj_%03d' % i] for i in range(60)]
    us_centers = fh['us_centers'].tolist()
    us_force_constants = fh['us_force_constants'].tolist()
    md_trajs = [fh['md_traj_%03d' % i] for i in range(20)]
    kT = float(fh['kT'])
```

Listing 1: downloading the data

2 The task

Use TRAM to construct an MEMM.
Find and plot the implied timescales and discretize the data into suitable metastable states. Plot the free energy of the system along the Na-CL distance axis. Find the free energy differences between the different metastable states. If you have time, compare the MEMM to an MSM which you construct from the unbiased data.

Note: you have been given both biased and unbiased trajectories. Consider how to use this information when discretizing your data and finding the implied timescales. It may help to visualize the biased and unbiased trajectories first.