Weighted Ensemble Simulation: Review of Methodology, Applications, and Software (2017)

▼ Table of Contents

ABSTRACT

- Benefits of WE
    - one of the first rare-event methods capable of yielding unbiased estimates of nonequilibrium observables.
    - superlinear scaling
    - unbiased observables: rate constants and equilibrium state populations
    - can work for atomistic to cellular scope
    - some applications include:
        - protein conformational transitions
        - (un)binding

        - assembly processes

INTRODUCTION

- routine to carry out simulations on the ms timescale
    - which can see protein conformational changes and binding processes.
- Standard MD is difficult to characterize kinetics of biological process events for the prohibitive computational cost.
- The emergence of more accurate FF (polarizable FF and QM/MM models) will truncate the efficiency.
- Path sampling can ↑ efficiency of simulating rare events
    - by shifting the focus to functional transition than stable state.
    - they exploit the fact that these transition states are fleeting and infrequent.
    - $t_b \ll t_{dwell}$ transition time magnitude less than stable or metastable region.
- Important perk is that no bias is introduced into the dynamics
    - like metadynamics, aMD, replica exchange
    - path sampling is rigorious and straightforward calculation of rate constants without assumptions
- The ability to efficiently provide kinetics observables and ensembles of unbiased pathway is highly complementary to state-of-the-art kinetics experiments.
    - Their perk is that these kinetic experiments can only measure overall rate constants of biological process
    - path sampling
        - can get rate constant of individual steps
        - provide insight into degree of diversity of pathways
        - yield ensembles of atomically detailed structures of transient states
            - not attainable from experiments

        - useful to identifying residues of possible kinetic significance that could be mutated to alter rate of biological process
- WE approach as a method is a splitting strategy.

THEORY

The Original Huber-Kim Algorithm

- Phase-space points are configurations
- Velocity is not part of the picture.
- For very large t,
    - distribution will relax to nonequilibrium if there is an addition or removal of particles, probability, or energy
    - distribution will relax to equilibrium if there isn't “”
- Resampling — converting one valid sample to another for the same distribution.
- The basic procedure of WE simulation
    1. initial configuration has been chosen that defines a set of trajectory walkers.
    2. Simulate all walkers for an arbitrary brief interval — same interval.
    3. resample trajectories statistically, maintaining the distribution of trajectories.
- Resampling trajectories is performed using bins
    - bins are regions into which configuration space has been subdivided.
    - WE targets maintaining a fixed number of walkers per bin, M.
- Resampling is done in two ways:
    - If # of walkers < M, each walker is replicated, or split, to create multiple identical copies of the trajectory.
        - The weight of all child trajectories must sum to weight of the parent.
        - Child weights are made equal, *usually*.
    - If # of walkers > M, then trajectories is pruned to equal # of walkers.

        - walkers are merged, and weights are summed.
- Each bin can have different targeted # of walkers.
- Choices to be made (system-specific):
    - bins are constructed
    - # of walkers are selected
    - resampling interval needs to be chosen

Equilibrium, Steady States, and Kinetics

- To calculate macroscopic rate constants between transitions without bias
    - $MFPT(A \rightarrow B) = \frac{1}{Flux(A \rightarrow B; SS')}$
    - Denominator is the probability per unit time entering state B in A-to-B steady state.

SOFTWARE

APPLICATION

- can yield continuous pathways and rate constants for any type of stochastic dynamics, MC, BD, and MD.
    - highlighting generality.
- Conformational sampling
    - attractive alternative to metadynamics, adaptive biasing force, replica exchange MD, particularly when calculation of kinetics observables is also of interest.
- Protein folding
- Protein-peptide binding
- Protein-ligand unbinding

CHALLENGES AND FUTURE DIRECTIONS

- Most probable path can be missed due to inadequate sampling of the slowest relevant motions.
- explore WE variation and reduce the correlations and hence improve sampling.