A simple method for automatically detecting equilibrated regions in molecular simulations

John D. Chodera[1,*]

[1]*Computational Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065*

(Dated: May 3, 2014)

Molecular simulations (molecular dynamics, Monte Carlo) are often initiated from configurations that are highly dissimilar to equilibrium samples, a practice which causes the appearance of a distinct initial transient in various mechanical observables computed over the timecourse of the simulation. Traditional practice in simulation data analysis recommends this initial transient portion be discarded as *equilibration*, but no simple, general automated procedure exists. Here, we consider a simple, automated, easy-to-implement procedure in which the final region of the simulation that maximizes the number of effectively uncorrelated samples is used. We present a reference implementation of this procedure and illustrate its application to synthetic and real data.

Keywords: molecular dynamics (MD); Monte Carlo (MC); Markov chain Monte Carlo (MCMC); equilibration; timeseries analysis; statistical inefficiency; integrated autocorrelation time

I. INTRODUCTION

* Corresponding author; john.chodera@choderalab.org