Pushing the timestep limit of molecular dynamics with hamiltonian monte carlo

Kyle A. Beauchamp+[1,*] and John D. Chodera[1,†]

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

(Dated: April 16, 2015)

The quantitative accuracy of molecular dynamics simulations is limited by timestep discretization error. This error can be eliminated by the use of metropolization, e.g. hamiltonian monte carlo. This rigorous approach has been largely unused by the molecular simulation community for reasons of interpretation and computational efficiency. Herein we combine multi-timestep integration, GPU accelerated molecular dynamics, and hamiltonian monte carlo to provide substantial speed improvements. Furthermore, the guaranteed thermodynamic fidelity provided by hamiltonian monte carlo enables the treatment of sampling as a blackbox optimization problem with little human intervention.

Keywords: molecular dynamics

I. INTRODUCTION

Molecular

II. THEORY

A. Quantifying Performance and Sampling

Quantifying sampling performance requires consideration of several distinct elements. The main objective is to draw uncorrelated samples from some target distribution $P(x)$.

1. Effective step size: $p_{accept} * \Delta t$ 2. Effective nanoseconds (simulation) per day (wall clock) 3. Effective sample size / ineffefiency

III. RESULTS

A. Hydrogen Mass Repartitioning

B. Choice of steps per HMC iteration

C. Multiple Timestep GHMC: MTSGHMC

D. Alanine Populations and escape times

E. XCGHMC and XCMTSGHMC

F. Raw Performance: The cost of fancy integrators

IV. CONCLUSIONS

Density

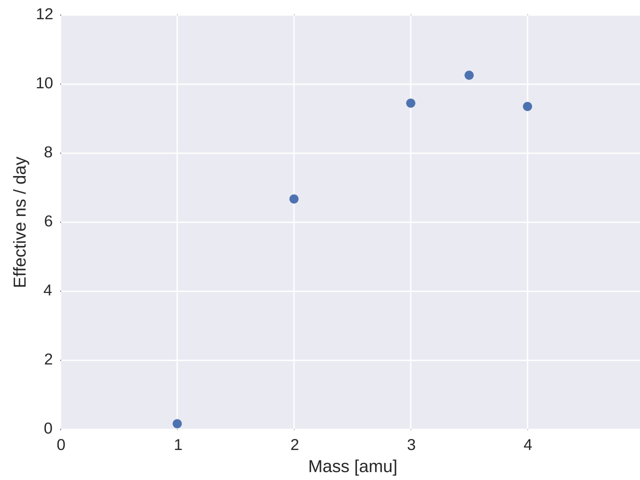

FIG. 1. HMR MASSES. The effect of hydrogen mass repartitioning on the effective performance, as measured by the raw performance times the acceptance rate.

	ns_per_day
name	
vvvr	90.7
ghmc1	33.7
ghmc10	82.7
vv	91.7
ghmcrespa20	69.8
verlet	103.7
ghmc20	90.7
langevin	103.9

TABLE I. Raw Performance of Various Integrators Raw performance on DHFR, assuming a 2 fs timestep. This does NOT incorporate the acceptance rate.

V. ACKNOWLEDGEMENTS

We thank Patrick B. Grinaway (MSKCC), Vijay S. Pande (Stanford University), Lee-Ping Wang (Stanford University), Peter Eastman (Stanford University), Robert McGibbon (Stanford University), Jason Swails (Rutgers University), David L. Mobley (University of California, Irvine), Christo-

* kyle.beauchamp@choderalab.org
† Corresponding author; john.chodera@choderalab.org

pher I. Bayly (OpenEye Software), Michael R. Shirts (University of Virginia), and members of Chodera lab for helpful discussions. Support for JMB was provided by the Tri-Institutional Training Program in Computational Biology and Medicine (via NIH training grant 1T32GM083937). KAB was supported in part by Starr Foundation grant I8-A8-058. JDC and KAB acknowledge partial support from NIH grant P30 CA008748. KAB, JLB, ASR, and JDC acknowledge the generous support of this research by the Sloan Kettering Institute.