Efficient Stochastic Simulation of Chemical Kinetics Networks Using A Weighted Ensemble Of Trajectories

Rory M. Donovan[1], Andrew J. Sedgewick[1], James R. Faeder[1] and Daniel M. Zuckerman[1]

We apply the "weighted ensemble" (WE) simulation strategy, previously employed in the context of molecular dynamics simulations, to a series of systems-biology models that range in complexity from a one-dimensional system to a system with 354 species and 3680 reactions. WE is relatively easy to implement, does not require extensive hand-tuning of parameters, does not depend on the details of the simulation algorithm, and can facilitate the simulation of extremely rare events in stochastic biochemical networks.

***Keywords* — Stochastic Dynamics, SSA, Signaling Networks, Rare Event Sampling, Weighted Ensemble.**

I. Introduction

STOCHASTIC behavior is an essential facet of biological processes such as gene expression, protein expression, and epigenetic modifications [1-2]. Stochastic chemical kinetics simulations are often used to study systems biology models of such processes [3]. One of the more common stochastic approaches, and the one employed in the present study, is the stochastic simulation algorithm (SSA), also known as the Gillespie algorithm.

As stochastic systems biology models approach the true complexity of the systems being modeled, it quickly becomes intractable to investigate rare behaviors using naïve ("brute-force") simulation approaches. By their very nature, rare events occur infrequently; confoundingly, they are often those of most interest. For example, the switching of a bistable system from one state to another may happen so infrequently that running a stochastic simulation long enough to see transitions is (extremely) computationally prohibitive [5]. This impediment only grows as model complexity increases, and as such it poses a serious hurdle for systems models as they grow more intricate.

II. Results

In addition to benchmarking on simple systems such as the enzymatic futile cycle and the Schlögl reactions, we applied weighted ensemble to a signaling model that is, to our knowledge, considerably more complex than any other biochemical system to which rare event sampling techniques have been applied. The reaction network in this model contains 354 chemical species and 3680 chemical reactions. This model describes association, dissociation, and phosphorylation reactions among four components: the receptor FcεRI, a bivalent ligand that aggregates receptors into dimers, and the protein tyrosine kinases Lyn and Syk. The weighted ensemble approach is able to characterize the probability distribution of phosphorylated receptors about six orders of magnitude more accurately in the tail of the distribution.

We also study a Cantor-Collins genetic switch. By setting up a feedback steady-state, and using the relation that the inverse of the probability flow into a state is equal to the mean first passage time to that state, weighted ensemble is able to find an estimate of the MFPT with greater precision than brute-force, using about one ten-thousandth the computational effort.

III. Conclusion

For the coupled stochastic reaction systems we study, WE is able to produce accurate and efficient approximations of the joint probability distribution for all chemical species for all time *t*. WE is also able to efficiently extract mean first passage times for the systems, via the construction of a steady-state condition with feedback. In all cases studied here, WE results agree with independent calculations, but significantly enhance the precision with which rare or slow processes can be characterized. Speedups over "brute-force" in sampling rare events via the Gillespie direct Stochastic Simulation Algorithm range from $\sim 10^{12}$ to $\sim 10^{18}$ for rare states in a distribution, and $\sim 10^{2}$ to $\sim 10^{4}$ for finding mean first passage times.

Acknowledgements: This work was funded by NSF grant MCB-1119091, NIH grant P41 GM103712, NIH grant T32 EB009403, and NSF Expeditions in Computing Grant (award 0926181).

[1]Department of Computational and Systems Biology, University of Pittsburgh. E-mail: donovanr@pitt.edu

References

[1] W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins, Nature 422, 633 (2003).
[2] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science (New York, N.Y.) 297, 1183 (2002)
[3] M. L. Blinov, J. R. Faeder, B. Goldstein, and W. S. Hlavacek, Bioinformatics (Oxford, England) 20, 3289 (2004).
[4] D. T. Gillespie, The Journal of Physical Chemistry 81, 2340 (1977).
[5] R. Allen, P. Warren, and P. ten Wolde, Physical Review Letters 94, 018104 (2005).