Automatic optimization of importance sampling parameters in weighted stochastic simulation algorithm

Min K. Roh[1,2], Bernie Daigle Jr.[1,3], Dan T. Gillespie[4], and Linda R. Petzold[5]

***Short Abstract* — The weighted stochastic simulation algorithm (wSSA) enables efficient estimation of the probability of a rare event in chemically reacting system within a specified time. The core of the wSSA is importance sampling, which is used to bias the system towards the desired rare event. However, the current algorithm requires the user to exercise insight and judgment in choosing the importance sampling parameters. We have developed a fully automated method based on information theory that chooses the optimal importance sampling parameters without any knowledge of the input system. We demonstrate its effectiveness on three different examples.**

***Keywords* — Weighted stochastic simulation algorithm, stochastic rare event simulation, importance sampling, information theory.**

I. BACKGROUND

THE stochastic simulation algorithm (SSA) and its variants are extensively used in the simulation of many biochemical systems. However, the use of SSA for computation of the probabilities of rare events is often not feasible because the required ensemble size to generate an accurate estimate may be prohibitive.

The weighted stochastic simulation algorithm [1] was recently developed by Kuwahara and Mura to alleviate the computational cost and improve accuracy. The primary component of the algorithm involves importance sampling (IS), which is used in the reaction selection procedure to bias the system toward the desired rare event. Gillespie *et al* [2] later extended the algorithm to include sample variance, which is used to compute the uncertainty of the estimate and find the optimal IS parameters. Currently, the IS parameters used by the wSSA are chosen by the user. However, this requires quite a bit of insight about the chemically reacting system and is not practical for complicated systems. A wSSA estimate using non-optimal IS parameters can be less accurate than the one from SSA.

We have developed a novel method that automatically optimizes a set of IS parameters without any knowledge of the input system. The method utilizes information theory to iteratively push the system toward the desired rare event and simultaneously optimize the IS parameters.

We tested the method on three different chemical systems: a reversible isomerization, an enzymatic futile cycle, and a modified yeast polarization model. We then calculated its computational gain, g, given by:

$$g \equiv \frac{n^{\mathrm{SSA}}}{n^{\mathrm{wSSA}}}, \tag{1}$$

where n^{SSA} and n^{wSSA} are the numbers of runs in each of the two methods needed to achieve comparable accuracy. We found that use of the optimized IS parameters yields a computational gain that ranges from two to four orders of magnitude. For the modified yeast polarization model, the computational gain suggests that the simulation time using the SSA to obtain an estimate of similar accuracy as the wSSA would be several days, wherein the wSSA simulation took only several minutes.

II. CONCLUSION

Manually choosing the values of IS parameters in wSSA becomes increasingly difficult as the system size grows. The initial set of parameters is usually formed by conjecture and is likely to yield a poor estimate. As the number of importance sampling parameters is equal to the total number of reactions, manual tuning of these parameters to produce an accurate estimate may not be computationally feasible for a large system. To overcome these difficulties, we have developed a fully automated method that optimizes a set of IS parameters. Consequently, our method may enable estimation of a rare event in large systems that have not been possible due to the time and computational constraint.

REFERENCES

[1] Kuwahara H, Mura I (2008) An efficient and exact stochastic simulation method to analyze rare events in biochemical systems. *J. Chem. Phys* **129**, 165101.

[2] Gillespie D, Roh M, Petzold L. (2009) Refining the weighted stochastic simulation algorithm. *J. Chem. Phys* **130**, 174103.

[1]These authors contributed equally to this work.

[2]Department of Computer Science, University of California Santa Barbara. E-mail: min@cs.ucsb.edu

[3]Department of Computer Science, University of California Santa Barbara. E-mail: bdaigle@gmail.com

[4]Dan T Gillespie Consulting, 30504 Cordoba Pl., Castaic, California 91384, USA

[5]Department of Computer Science, University of California Santa Barbara.