Reachability Analysis for Stochastic Hybrid Systems Using Multilevel Splitting

Derek Riley[1], Xenofon Koutsoukos[1], and Kasandra Riley[2]

[1] ISIS/EECS,
Vanderbilt University, Nashville, TN 37235, USA
Derek.Riley,Xenofon.Koutsoukos@vanderbilt.edu
[2] Howard Hughes Medical Institute
Yale University, New Haven, CT, USA
Kasandra.Riley@yale.edu

1 Introduction

Biomedical research is increasingly using formal modeling and analysis methods to improve the understanding of complex systems. Verification methods for Stochastic Hybrid Systems (SHSs) are burdened with the curse of dimensionality; however, probabilistic analysis methods such as Monte Carlo (MC) methods can be used to analyze larger systems. MC methods are useful for estimating probabilities of event occurrences in SHS, but large and complex systems may require prohibitively large computation time to generate sufficient accuracy. In this work we present the multilevel splitting (MLS) variance reduction technique that has the potential to reduce the variance of MC methods by an order of magnitude significantly improving both their efficiency and accuracy [1].

This work presents an implementation of MLS methods for safety analysis of SHS. We apply the approach for safety analysis of the glycolysis process, which we model with a SHS model with two discrete states and 22 continuous variables. We also present experimental data along with accuracy and efficiency analysis. Further, the technique is parallelized to increase the efficiency, and we present the scalability of the parallelization.

2 SHS Model of Glycolysis

Glycolysis is a series of biochemical reactions that converts carbohydrates into chemicals and energy in a currency useful to cells. As it is a fundamental process to all living cells, it has been studied and modeled extensively in many organisms. Although the individual steps of glycolysis have been thoroughly examined, the interaction of glycolytic enzymes, substrates, and products with the intracellular environment is not fully understood. Modeling and simulating glycolysis using SHS can further our understanding of contextual cellular respiration.

Twenty-two chemical species and 37 chemical reactions have been identified which play an important role in glycolysis. The reaction rates for the system have been developed in previous work and can be found in [2]. The model presented

R. Majumdar and P. Tabuada (Eds.): HSCC 2009, LNCS 5469, pp. 460–464, 2009.

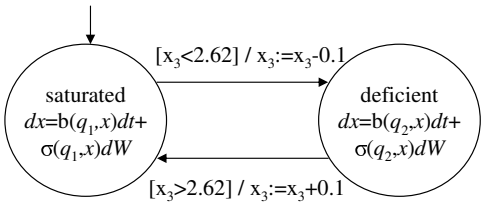

Fig. 1. SHS model of glycolysis

in [2] is a deterministic model, but the chemical reactions in the real system actually behave in a probabilistic manner due to the uncertainty of molecular motion, so we have developed a stochastic model in a similar manner as was done in [3].

We have added discrete dynamics to the original glycolysis model to capture the concept of 'feeding' the yeast. Glucose must be added to the system to continue production of the energy molecules, and when the concentration of glucose diminishes, the amount of energy molecules that the system can produce decreases. In many organisms, this reduction in energy output triggers mechanisms which encourage the introduction of more glucose (i.e. feeding). Therefore, we have modeled this behavior using a SHS with two states: saturated and deficient. In the saturated state, the glucose intake is fairly low, and in the deficient state, the glucose intake is much higher. Switching between the states is regulated by the concentration of ATP (x_3). A probabilistic reset map is used on the transition to avoid Zeno behavior. Figure 1 depicts the graphical version of the SHS model.

It is important for the cell performing glycolysis to maintain a certain concentration of Glucose x_1 to maintain cell health. Therefore, we want to determine if the state trajectories will avoid the set $U = \{(q, x) : x_1 < 2.5\}$.

3 Multilevel Splitting

We denote $s(t)$ the SHS trajectory, τ_{max} the maximum simulation time, and we define the stopping time $\tau_U = inf\{t > 0 : s(t) \in U\}$. We want to compute the probability that a trajectory will hit the unsafe set $P_{hit} = \mathbb{P}[\tau_U < \tau_{max}]$. P_{hit} can be estimated using MC methods; however, they are often computationally too expensive to generate estimators with small variance. MLS is an adaptation of MC methods that reduces the overall variance of the estimator by increasing the density of simulation trajectories near U [1].

MLS trajectories use importance values v_i to represent the amount of influence a trajectory has on P_{hit}. Initially $v_i = 1/n$ where n is the total number of trajectories. We define splitting levels using proper supersets of the unsafe set U: $U \subset U_1 \subset U_2 \subset ... \subset U_g$. When a trajectory crosses from a bigger set U_k into a smaller set U_{k-1}, the trajectory is split into j new trajectories, the importance value of the current trajectory is split between the new forked trajectories, and the total number of trajectories n_m is incremented by $j - 1$.

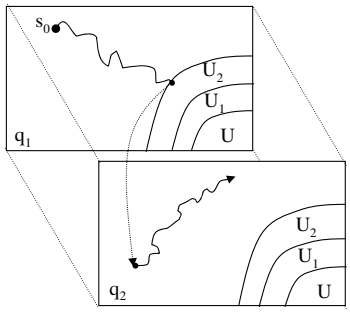

Fig. 2. MLS problem in a hybrid state space

The probability of hitting the unsafe set U is estimated for MLS methods by $\widehat{P_{hit}} = \sum_{i=1}^{n_m} H_i v_i$ where $H_i = 1$ if the trajectory eventually reaches U and $H_i = 0$ otherwise. Even though the resulting split trajectories are not completely independent, $\widehat{P_{hit}}$ is an unbiased estimator. The efficiency and accuracy of the estimator are dictated by the boundary placement, splitting policy, and dynamics of the model. The efficiency can be evaluated using $Eff\left[\widehat{P_{hit}}\right] = \frac{1}{Var \cdot C}$ where C is the expected execution time to compute the estimator, and Var is the variance of the estimator [1].

The discrete boundaries and reset maps present in SHS can create discontinuities which can cause accuracy and efficiency challenges for variance reduction methods. Figure 2 demonstrates a MLS scenario for a SHS where the trajectory crosses a splitting and hybrid boundary simultaneously. A hybrid trajectory starts at state $s_0 = (q_1, x_0)$, and evolves until it reaches the boundary for U_2 or the guards for a hybrid transition are satisfied. In this scenario, both the hybrid transition is fired and the splitting level is crossed, and the reset of the hybrid transition updates the state of the trajectory to $s = (q_2, x_t)$. Because the new state is not in the splitting region U_2, splitting the trajectory before applying the reset will not necessarily reduce the variance, and will decrease the efficiency, so it should be avoided. This problem is further exacerbated if the number of splits at a level is large because poor splitting choices decrease efficiency without increasing accuracy.

We have implemented the MLS algorithm for SHS using the simulation methods described in [4] to generate accurate and efficient SHS trajectories. Our MLS implementation ensures that discrete transitions are fired before testing splitting boundaries to avoid the potential efficiency loss of the boundary problem shown in Figure 2. The number of required simulations to achieve a sufficiently small variance may still be quite large even when using our algorithm, so we use parallel methods to improve overall efficiency. There are no dependencies between MLS simulations, so trajectories can be parallelized by running simulations concurrently on multiple processors. This type of parallelization has been used previously with MC methods [5], and care must be taken to ensure that the random number generators do not introduce bias.

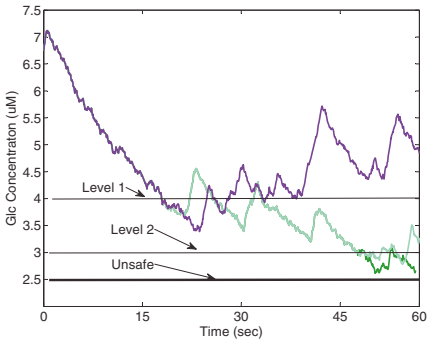

Fig. 3. Trajectory of the glycolysis model with MLS

Table 1. Performance and variance results

Simulation	Estimator	Var	$Time(C)$	Eff	P_{hit}
$sim1$	MC	623	398	0.0000040	0.12
$sim2$	MC	625	500	0.0000032	0.0625
$sim1$	MLS	711	128	0.0000110	0.0625
$sim2$	MLS	691	123	0.0000118	0.0625

Table 2. Parallel performance results

Processors	Time to Execute (m)
16	8.6
8	8.5
4	8.2
2	8.3
1	8.5

4 Experimental Results

Single trajectories of a model can be used to collect specific information about the system. In Figure 3, we show a single trajectory of the glycolysis model using MLS with two levels and two splits at each level.

To evaluate the efficiency of our methods for the glycolysis model, we used 16 trajectories for the MC methods and 2 MLS trajectories. This allows both methods to reach the same potential accuracy because the MLS scenario used three levels with two splits at each level yielding $2^4 = 16$ potential forked trajectories. In Table 1 we compare the variance and execution times using an order 0.5 simulation method ($sim1$) and an order 1.0 method with probabilistic boundary detection ($sim2$) from [4] with both MC methods and MLS methods. The data shows a significant efficiency improvement for the MLS estimator without a significant decrease in variance.

We also performed experiments to test the parallel scalability of our algorithm. We found that the MLS algorithm took virtually the same amount of

time regardless of the number of processors it was run on as seen in Table 2. The Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University provides the parallel computing resources for our experiments (www.accre.vanderbilt.edu).

5 Conclusions and Future Work

Analysis of SHS using Monte Carlo methods with variance reduction is an important technique which has the potential to expose insights into complex models efficiently. The SHS analysis method we present in this work demonstrates an efficient, accurate variance reduction method with parallelization, but it requires significant domain knowledge to determine appropriate splitting parameters. The method holds promise to provide further analysis capability for SHS simulation methods as well. In the future we will be investigating methods for selecting boundary placement and splitting policies based on methods presented for other splitting techniques. Our goal is to find an optimal policy for selecting the MLS parameters.

Acknowledgements. This research is partially supported by the National Science Foundation (NSF) CAREER grant CNS-0347440.

References

1. L'Ecuyer, P., Tuffin, B.: Splitting for rare-event simulation. In: Winter Simulation Conference, pp. 137–148 (2006)
2. Hynne, F., Dano, S., Sorensen, P.: Full-scale model of glycolysis in Saccharomyces cerevisiae. Biophysical Chemistry 94, 121–163 (2001)
3. Riley, D., Koutsoukos, X., Riley, K.: Modeling and simulation of biochemical processes using stochastic hybrid systems: The sugar cataract development process. In: Egerstedt, M., Mishra, B. (eds.) HSCC 2008. LNCS, vol. 4981, pp. 429–442. Springer, Heidelberg (2008)
4. Riley, D., Koutsoukos, X., Riley, K.: Simulation of stochastic hybrid systems with switching and reflective boundaries. In: Winter Simulation Conference, pp. 804–812 (2008)
5. Troyer, M., Ammon, B., Heeb, E.: Parallel object oriented monte carlo simulations. In: Caromel, D., Oldehoeft, R.R., Tholburn, M. (eds.) ISCOPE 1998. LNCS, vol. 1505, pp. 191–198. Springer, Heidelberg (1998)