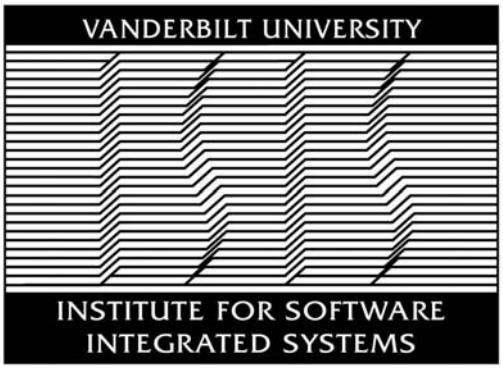

Institute for Software Integrated Systems
Vanderbilt University
Nashville Tennessee 37235

TECHNICAL REPORT

TR #: ISIS-06-708
Title: Safety Analysis of Sugar Cataract Development Using Stochastic Hybrid Systems
Authors: Derek Riley, Xenofon Koutsoukos, Kasandra Riley

Safety Analysis of Sugar Cataract Development Using Stochastic Hybrid Systems

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

[1] Institute for Software Integrated Systems
Department of Electrical Engineering and Computer Science
Vanderbilt University
Nashville, TN 37235, USA
Derek.Riley, Xenofon.Koutsoukos@vanderbilt.edu
[2] Mayo Foundation
Guggenheim 1611-C
200 First Street SW
Rochester, MN 55905, USA
riley.kasandra@mayo.edu

Abstract. Modeling and analysis of biochemical systems are critical problems because they can provide new insights into systems which can not be easily tested with real experiments. One such biochemical process is the formation of sugar cataracts in the lens of an eye. Analyzing the sugar cataract development process is a challenging problem due to the highly-coupled chemical reactions that are involved. In this paper we model sugar cataract development as a stochastic hybrid system. Based on this model, we present a probabilistic verification method for computing the probability of sugar cataract formation for different chemical concentrations. Our analysis can potentially provide useful insights into the complicated dynamics of the process and assist in focusing experiments on specific regions of concentrations. The verification method employs dynamic programming based on a discretization of the state space and therefore suffers from the curse of dimensionality. To verify the sugar cataract development process we have developed a parallel dynamic programming implementation that can handle large systems. Although scalability is a limiting factor, this work demonstrates that the technique is feasible for realistic biochemical systems.

1 Introduction

Modeling and analysis of biochemical systems are important tasks because they can unlock insights into the complicated dynamics of systems which are difficult or expensive to test experimentally. A variety of techniques have been used to model biochemical systems, but the effectiveness of the analysis techniques is often limited by tradeoffs imposed by the modeling paradigms. Stochastic differential equations have been used to model biochemical reactions [13, 3]; however, analysis of these models has mainly been limited to simulation. Hybrid systems have also been used to model biochemical systems [1, 12]; however, verification

methods based on deterministic hybrid systems fail to capture the probabilistic nature of some biochemical processes and therefore may not be able to correctly analyze certain systems. Stochastic Hybrid Systems (SHS) have been used to capture the stochastic nature of biochemical systems but have previously only been used for simulations [31] or analysis of systems with simplified continuous dynamics [16].

In this paper we analyze the biochemical process of sugar cataract development in the lens of an eye. The enzyme sorbitol dehydrogenase catalyzes a reversible oxidation of sorbitol and other corresponding keto-sugars. An accumulation of sorbitol in the eye is theorized to be the main factor in the development of a sugar cataract. The chemical reactions and kinetic constants for the model have been previously studied [26]. Understanding the exact conditions that lead to the development of sugar cataracts will help scientists better predict and prevent the condition [3]. Our analysis results can potentially provide useful insights into its complicated dynamics and assist in focusing experiments.

We model the sugar cataract development problem using SHS and use a dynamic programming verification method based on a discretization of the state space [21]. The proposed method suffers from the curse of dimensionality. Therefore, we have developed a parallel dynamic programming implementation of the verification algorithm that can handle large systems. Although scalability is a limiting factor, this work demonstrates that the technique is feasible for realistic biochemical systems.

The organization for the rest of the paper is as follows: Section 2 describes the related work, Section 3 describes modeling of biochemical systems and the sugar cataract development process using SHS, Section 4 describes the probabilistic verification method, Section 5 presents our experimental results, and Section 6 concludes the work.

2 Related Work

Many systems in the biological sciences can benefit from formal modeling and analysis methods. A variety of modeling techniques have been used to model species population evolution to molecular dynamics [19]. As computing power has increased, modeling and simulation approaches have evolved to take advantage of increased computational power to improve accuracy and speed. Stochastic Differential Equations (SDE) have been used for modeling cell signaling pathways and molecular motion [12, 23, 3]. Since only specialized cases of SDEs can be solved analytically, the vast majority of models are simulated using Monte Carlo techniques. The original cell signaling models were simulated using a fixed-step stochastic simulation algorithm presented in [13]. The simulation algorithm was fairly inefficient for large models, so computational improvements have been made since [3].

Hybrid systems have been used for modeling biological systems in order to capture the complicated dynamics using well-defined abstractions. Biomolecular network modeling is accomplished by using differential equations to model

feedback mechanisms and discrete switches to model changes in the underlying dynamics [1]. Biological protein regulatory networks have been modeled with hybrid systems using linear differential equations to describe the changes in protein concentrations and discrete switches to activate or deactivate the continuous dynamics based on protein thresholds [12].

Stochastic hybrid systems further improve on the benefits of hybrid systems by providing a more realistic probabilistic framework for modeling real-world biochemical systems. A modeling technique that uses SHS to construct models for chemical reactions involving a single reactant specie is presented in [16]. A genetic regulatory network was modeled with a SHS model and compared to a deterministic model in [18]. SHS models of biochemical systems have been developed and simulated using hybrid simulation algorithms in [14, 31].

This paper adopts a SHS model that is a special case of the general model presented in [7]. Related models have been presented in [15] with the emphasis on modeling and analysis of communication networks and in [5] for simulation of concurrent systems. SHS can be viewed as an extension of piecewise-deterministic processes [10] that incorporate stochastic continuous dynamics. Reachability of such systems has been studied in [8].

Reachability properties for continuous and hybrid systems have been characterized as viscosity solutions of variants of HJB equations in [25, 27]. Extensions of this approach to SHS and a toolbox based on level set methods have been presented in [28]. A technique for probabilistic reachability verification for discrete time SHS based on the interpretation of the safety verification problem as an optimal control problem for a certain controlled Markov process has been presented in [2].

This paper employs a reachability analysis method based on discrete approximations. Discrete approximation methods based on finite differences have been studied extensively in [24] and the references therein. Based on discrete approximations, the reachability problem can be solved using algorithms for discrete processes [30]. The approach has been applied for optimal control of SHS given a discounted cost criterion in [20]. For verification, the discount term cannot be used and convergence of the value function can be ensured only for appropriate initial conditions. A related grid based method for safety analysis of stochastic systems with applications to air traffic management has been presented in [17]. Our approach is similar but using viscosity solutions we show the convergence of the discrete approximation methods.

Reachability analysis for SHS can also be accomplished using Monte Carlo methods. Multiple stochastic simulations are used to determine the reachability probability for an initial state of a SHS. Confidence intervals and accuracy probabilities can be selected by adjusting the number of simulations [29]. Stochastic π-calculus is another effective modeling framework for biological processes. The models generated by these techniques are continuous time Markov Chains so verification techniques and tools can be used to analyze the systems [23]. The modeling used by π-calculus is different from that used in SHS, and comparison of these methods is a subject of future work.

3 Modeling Biochemical Reactions using SHS

3.1 Dynamics of Biochemical Reactions

All cellular function of living organisms is governed by complex systems of coupled biochemical reactions. A reaction specifies all chemical species which react (reactants) and are produced (products). A kinetic constant k, associated with each reaction, numerically describes the affinity for the reactants to produce the products in defined temperature and pressure conditions.

Experimental analysis is used to physically measure the variation in individual concentrations of the chemical species in a biochemical system. However, understanding the dynamical behavior of biochemical systems requires running many experiments that can be time consuming, tedious, unsafe, or costly. Developing and analyzing dynamical models for capturing the evolution of individual chemical species concentrations can reduce the number of experiments needed.

Discrete models are a natural modeling paradigm for biochemical systems because reactions can be considered as occurring at specific points in time, and when a reaction occurs, individual molecules interact and produce new molecules. Discrete models update the concentrations of the involved reactants and products at a certain reaction rate based on the stoichiometry defined by the reaction.

Chemical reactions are inherently probabilistic because of the unpredictability of molecular motion [11], so their dynamics are best described by stochastic models. Discrete stochastic models of reactions can be created by describing a reaction j as firing at a rate a_j [9]. When the reaction fires, the concentrations of the reactants and products are reset to the appropriate updated values. Table 1 shows the rates and resets for several examples of different types of reactions. For example, when the reaction $X \rightarrow Z$ occurs, a molecule of X is consumed and a molecule of Z is produced denoted by $x- = 1$ and $z+ = 1$ respectively where x and z are the quantities of molecules of chemical species X and Z, and k_i is the kinetic constant for reaction i.

Reaction	a_j	Reset
$X \rightarrow Z$	$k_1 x$	$x- = 1;\ z+ = 1;$
$X + Y \rightarrow 2Z$	$k_2 xy$	$x- = 1;\ y- = 1\ ;\ z+ = 2;$
$2X \rightarrow Z$	$1/2 * k_3 x(x-1)$	$x- = 2;\ z+ = 1;$
$2X + Y \rightarrow 2Z$	$1/2 * k_4 x(x-1)y$	$x- = 2;\ y- = 1;\ z+ = 2;$
$3X \rightarrow Z$	$1/6 * k_5 x(x-1)(x-2)$	$x- = 3;\ z+ = 1;$

Table 1. Example Reaction Rates and Resets

Reactions occur at different speeds depending on the concentrations of chemicals and the kinetic constant for each reaction. "Slow" reactions occur when reaction rates and concentrations are small enough and they can be modeled and simulated efficiently using discrete stochastic techniques. However, discrete simulations become inefficient when there are large concentrations of molecules

and/or fast reaction rates. In such cases the reaction will occur very frequently and the discrete simulation will need to consider a large number of transitions in a short period of time. "Fast" reactions occur at a rate that is fast enough or in high enough concentrations to consider as occurring at a constant rate. Such reactions can be modeled more efficiently as stochastic continuous models assuming the reactions happen in a well-mixed solution [31].

The rate of change of each chemical species is calculated using the chemical dynamics from the biochemical reactions. Suppose that we have a system of M chemical reactions and N chemical species. We define x_i as the concentration of the ith chemical species in micro-Molarity (μM), M_{fast} as the number of fast reactions, a_j as the reaction propensity of the jth reaction, and W as an M_{fast}-dimensional Wiener process. The stoichiometric matrix v is a (M_{fast} X N) matrix which holds values representing the concentration of chemical species lost or gained in each reaction. The following equation describes the dynamics for each of the i chemical species [31].

$$dx_i = \sum_{j=1}^{M_{fast}} v_{ji} a_j(x(t))dt + \sum_{j=1}^{M_{fast}} v_{ji} \sqrt{a_j(x(t))} dW_j \tag{1}$$

Discrete and continuous models consider only slow or only fast chemical reactions, but real biochemical systems often contain a mixture of both fast and slow reactions. In a such a situation it is most efficient to use a hybrid modeling approach to take advantage of the efficiency of continuous modeling while still keeping the accuracy of discrete modeling [31].

Stochastic hybrid systems are ideal for modeling biochemical systems with both fast and slow chemical reactions systems because they are able to model continuous and discrete dynamics in a stochastic framework. Fast reactions are modeled using the continuous stochastic dynamics techniques presented earlier, and slow reactions are modeled as discrete transitions with probabilistic rates and resets.

To determine which reactions are fast or slow, one must analyze the rates using the kinetic parameters and quantities of each reactant involved. The reaction rate range can be determined by analyzing the rate a_j from Table 1 over the entire range of possible chemical concentrations. To determine the smallest rate, the smallest concentrations for each chemical species should be used. Similarly, the largest rate can be determined by using the highest concentrations in the range. Reactions can be considered slow if the reaction rate never exceeds 100 reactions per second, otherwise reactions can be modeled as fast reactions. If a reaction has a range that spans 100 reactions per second, the reaction can be classified as either fast or slow.

3.2 Sugar Cataract Development

A sugar cataract is a type of cataract which distorts the light passing through the lens of an eye by attracting water to the lens when an excess of sorbitol is present. Often these cataracts are formed in the eyes of diabetes patients who

do have highly fluctuating blood sugar levels. Several factors affect the accumulation of sorbitol including the amount of the enzyme sorbitol dehydrogenase. Sorbitol dehydrogenase catalyzes the reversible oxidation of sorbitol and other polyalcohols to the corresponding keto-sugars [26].

The chemical species and concentration ranges for the sugar cataract development process for bovine lens are described in Table 2. The bovine lens data is used as a standard model for human cataract development. The ranges are bounded and are estimated using realistic concentration values derived from experimental data and Michaelis-Menten constants (Km) defined as the rate of the reaction at half-maximal velocity [26]. Table 3 describes the seven reactions and rates involved in sugar cataract development. The rate is calculated based on the concentrations presented in Table 2 and the kinetic constants presented in Table 3.

Reactant	Variable	Min conc. (μM)	Max conc. (μM)
$NADH$	x_1	0.0005	20.0005
$E-NADH$	x_2	0.0005	20.0005
NAD^+	x_3	0.0009	10.0009
$E-NAD^+$	x_4	0.0009	10.0009
sorbitol dehydrogenase (E)	x_5	0.0002	0.2002
fructose (F)	x_6	0.2	2000.2
sorbitol (S)	x_7	0.2	2000.2
Inactive form of E (Z)	-	0.000002	0.200002

Table 2. Chemical species properties for the sugar cataract model

Reaction	Kinetic constant	Rate
$E+NADH \rightarrow E-NADH$	$k_1 = 6.2$	Fast
$E-NADH \rightarrow E+NADH$	$k_2 = 33$	Fast
$E-NADH+F \rightarrow E-NAD^+ + S$	$k_3 = 0.0022$	Fast
$E-NAD^+ + S \rightarrow E-NADH+F$	$k_4 = 0.0079$	Fast
$E-NAD^+ \rightarrow E+NAD^+$	$k_5 = 227$	Fast
$E+NAD^+ \rightarrow E-NAD^+$	$k_6 = .61$	Fast
$E \rightarrow Z$	$k_7 = 0.0019$	Slow

Table 3. Sugar cataract reactions and kinetic constants

The slow reaction $E \rightarrow Z$ describes the conversion of the enzyme (E) into its inactive form (Z) at a rate of $k_7 x_5$ according to Table 1. When the reaction occurs, the number of molecules of E will be decreased by one and the concentration will be decreased by $d = 10^{-21}$ μ Molar.

Each of the six fast reactions are modeled using the SDE (1). The inactive form of E (Z) is not a reactant in any of the chemical equations, so its concentration is not modeled. The equations describe the rates of change of the individual chemical species and are given below.

$$
\begin{aligned}
dx_1 &= (-k_1x_1x_5 + k_2x_2)dt - \sqrt{k_1x_1x_5}dW_1 + \sqrt{k_2x_2}dW_2 \\
dx_2 &= (k_1x_1x_5 - k_2x_2 - k_3x_2x_6 + k_4x_4x_7)dt + \\
&\quad \sqrt{k_1x_1x_5}dW_1 - \sqrt{k_2x_2}dW_2 - \sqrt{k_3x_2x_6}dW_3 + \sqrt{k_4x_4x_7}dW_4 \\
dx_3 &= (k_5x_4 - k_6x_3x_5)dt + \sqrt{k_5x_4}dW_5 - \sqrt{k_6x_3x_5}dW_6 \\
dx_4 &= (k_3x_2x_6 - k_4x_4x_7 - k_5x_4 + k_6x_3x_5)dt + \\
&\quad \sqrt{k_3x_2x_6}dW_3 - \sqrt{k_4x_4x_7}dW_4 - \sqrt{k_5x_4}dW_5 + \sqrt{k_6x_3x_5}dW_6 \\
dx_5 &= (-k_1x_1x_5 + k_2x_2 + k_5x_4 - k_6x_3x_5)dt - \\
&\quad \sqrt{k_1x_1x_5}dW_1 + \sqrt{k_2x_2}dW_2 + \sqrt{k_5x_4}dW_5 - \sqrt{k_6x_3x_5}dW_6 \\
dx_6 &= (-k_3x_2x_6 + k_4x_4x_7)dt - \sqrt{k_3x_2x_6}dW_3 + \sqrt{k_4x_4x_7}dW_4 \\
dx_7 &= (k_3x_2x_6 - k_4x_4x_7)dt + \sqrt{k_3x_2x_6}dW_3 - \sqrt{k_4x_4x_7}dW_4
\end{aligned}
$$

Biologists have determined that a ratio of sorbitol to fructose that is greater than one is correlated to the beginning stages of sugar cataract formation [4]. It has been shown that fructose (x_6) and sorbitol dehydrogenase (x_5) play a significant role in the accumulation of sorbitol (x_7) in the eye which in turn begins the formation of sugar cataracts.

3.3 SHS Model of Sugar Cataract Development

This section describes a formal SHS model for the sugar cataract development following the formalism presented in [21]. Let $x = [x_1, \ldots, x_7]^T$ denote the continuous state taking values in $X \subset \mathbb{R}^7$ where the set X is defined by the concentration ranges in Table 2. The continuous dynamics of the SHS are given by the SDE

$$dx = b(x)dt + \sigma(x)dW \tag{2}$$

describing the dynamics of the fast reactions for the sugar cataract development presented in Subsection 3.2 where $b : \mathbb{R}^7 \to \mathbb{R}^7$ is the drift vector, $\sigma(x) : \mathbb{R}^7 \to \mathbb{R}^{7\times 6}$ is the dispersion matrix, and $W(t)$ is an $\mathbb{R}^6$-valued Wiener process.

To capture the discrete dynamics due to the slow chemical reaction, it is sufficient to consider a hybrid system with one discrete state and with a self-transition representing an occurrence of the slow reaction. When the discrete transition occurs, the concentration of E (x_5) jumps instanteneously according to the assignment $x_5 := x_5 - d$.

Let $Q = \{q\}$ denote the discrete state set and $\lambda(x) = k_7x_5$ the transition rate function that is associated with the discrete transition. To represent the state jumps, we define a reset map $R : \mathbb{R}^n \times \mathbb{R}^n \to \{0, 1\}$ by

$$R(x, x') = \begin{cases} 1 \text{ if } x'_5 = x_5 - d \\ 0 \text{ otherwise} \end{cases}.$$

The SHS is then defined as $(Q, \mathbb{R}^7, b, \sigma, Init, \lambda, R)$ where $Init$ is the initial condition for the concentrations of the reactants.

Between transitions, the continuous state evolves according to the SDE where the solution is understood using the Itô stochastic integral. The occurrence of a discrete transition is governed by an exponential distribution characterized by the state-dependent transition rate $\lambda(x)$ and, upon occurrence of a transition, the continuous state x is reset according the the reset map $R(x, x')$.

The functions $b(x)$ and $\sigma(x)$ are bounded and Lipschitz continuous in $x \in X$ and thus the SDE has a unique solution. The transition rate function $\lambda(x) = k_7 x_5$ is a bounded and measurable function which is integrable for every sample path x_t, and therefore, the expected value of the number of discrete transitions in a finite interval $[0, t]$ will be finite.

As described in Table 2, the concentrations of the sugar cataract development system are assumed to be bounded. Further, since all concentrations are positive, it is reasonable to assume that the diffusion term for the SDE is non-degenerate, i.e. $a(x) = \sigma(x)\sigma^T(x)$ is positive definite for every $x \in X$. Given these assumptions, the SHS for the sugar cataract development is a special of the SHS model described in [21]. In particular, this model has one discrete state and one discrete transition with a deterministic reset map.

For the sugar cataract development, a ratio of sorbitol to fructose that is greater than one is correlated to the beginning stages of the sugar cataract formation [3]. Therefore, we can define the set of safe states as the set of all concentrations that satisfy $x_7 - x_6 < 1$ and apply the reachability analysis method presented in [21].

4 Probabilistic Verification

4.1 Reachability Analysis

Given the set of safe states $B = \{x \in X : x_7 - x_6 < 1\}$, we consider the verification problem of computing the probability that the system execution from an arbitrary (safe) initial state will exit the safe set indicating the beginning stages of sugar cataract development. We denote ∂B and $\bar{B} = B \cup \partial B$ the boundary and the completion of B respectively. Consider the stopping time $\tau = \inf\{t \geq 0 : s(\tau^-) \in \partial B\}$ which is the first hitting of the boundary ∂B. Let x be an initial state in B, then we define the function $V : \bar{B} \to \mathbb{R}$ by

$$V(x) = \begin{cases} E_x[I_{(x(\tau^-) \in \partial B)}], & x \in B \\ 1, & x \in \partial B \end{cases}.$$

The function $V(x)$ can be interpreted as the probability that a trajectory starting at x will reach the boundary ∂B of the safe set, i.e. the probability that the system is unsafe and sugar cataract formation may begin.

The value function V that characterizes the safety of sugar cataract formation can be described as the viscosity solution of a Hamilton-Jacobi-Bellman (HJB) equation. This function is similar to the value function for the exit problem of a

standard stochastic diffusion, but the running and terminal costs depend on the function itself. This dependence captures the effects of the discrete dynamics to the value function for the exit problem.

First, assuming that $B \subset X$, we define a bounded function $c : \bar{B} \to \mathbb{R}_+$ continuous in x such that

$$c(x) = \begin{cases} 1, & \text{if } x \in \partial B \\ 0, & \text{otherwise} \end{cases} .$$

The next proposition presents the HJB equation for the problem. The proof is a straightforward application of the results presented in [21] to the SHS of the sugar cataract development.

Proposition 1. *Define* $L^V(x) = \lambda(x)V(y)R(x,y)$ *and* $\psi^V(x) = c(x)+V(y)R(x,y)$. *Then,* V *is the unique viscosity solution of the equation*

$$\mathcal{H}_V\left(x, V, D_x V, D_x^2 V\right) = 0 \text{ in } B$$

with boundary conditions

$$V(q,x) = \psi^V(q,x) \text{ on } \partial B$$

where

$$\mathcal{H}_V\left(x, V, D_x V, D_x^2 V\right) = b(x)D_x V + \frac{1}{2} tr(a(x)D_x^2 V) + \lambda(x)V + L^V(x).$$

4.2 Numerical Methods Based on Dynamic Programming

One of the advantages of characterizing reachability properties using viscosity solutions is that for computational purposes we can employ numerical algorithms based on discrete approximations. We use an approximation method based on finite differences and we present an iterative algorithm based on dynamic programming for computing the solution. The main characteristic of the approach is that the solution based on the discrete approximations converges to the one for the original stochastic hybrid system as the discretization becomes finer.

We employ the finite difference method presented in [24] to compute locally consistent Markov chains (MCs) that approximate the SHS while preserving local mean and variance. We consider a discretization of the state space characterized by the approximation parameter $h > 0$ representing the distance between neighboring points. By abuse of notation, we denote the sets of boundary and interior points in the state space X as ∂X^h and X^h respectively. By the boundness assumption, the approximating MC will have finitely many states which are denoted by $\xi_n^h,\ n = 1, 2, \ldots, N$.

Consider the continuous evolution of the SHS between jumps and assume that the state is x. The local mean and variance given the SDE (2) on the interval $[0, \delta]$ are

$$E[x(\delta) - x] = b(x(t))\delta + o(\delta)$$
$$E[(x(\delta) - x)(x(\delta) - x)^T] = a(x(t))\delta + o(\delta).$$

Let $\{\xi_n^h\}$ describe the MC on X^h with transition probabilities denoted by $p_D^h(x, x')$. A locally consistent MC must satisfy

$$E[\Delta\xi_n^h] = b(x)\Delta t^h(x) + o(\Delta t^h(x))$$

$$E[(\Delta\xi_n^h - E[\Delta\xi_n^h])(\Delta\xi_n^h - E[\Delta\xi_n^h])^T] = a(x(t))\Delta t^h(x) + o(\Delta t^h(x))$$

where $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h, \xi_n^h = x$ and $\Delta t^h(x)$ are appropriate interpolation intervals (or the "holding times") for the MC.

The diffusion transition probabilities $p_D^h(x, x')$ and the interpolation intervals can be computed systematically from the parameters of the SDE (details can be found in [24]). To incorporate the effect of the transition rate in the approximating MC, consider the jumps with transition rate $\lambda(x)$ and reset map $R(x, x')$. Suppose that the state has just changed $\xi_n^h = x$. The probability that a jump will occur on $[t, t + \delta)$ conditioned on the past data can be approximated by

$$P[x \text{ jumps on } [t, t + \delta)|x(s), w(s), s \leq t] = \lambda(x)\delta + o(\delta).$$

Therefore, with probability $1 - \lambda(x)\Delta t^h(x) - o(\Delta t^h(x))$ the next state is determined by the diffusion probabilities p_D^h and with probability $\lambda(x)\Delta t^h(x) + o(\Delta t^h(x))$ there is a jump and the next state is determined by the reset map $R(x, x')$. Therefore, the transition probabilities are defined by

$$\begin{aligned} p^h(x, x') = (1 - \lambda(x)\Delta t^h(x) - o(\Delta t^h(x)))p_D^h(x, x') + \\ (\lambda(x)\Delta t^h(x) + o(\Delta t^h(x)))R(x, x') \end{aligned} \tag{3}$$

Let $\bar{B}^h = \bar{X}^h \cap \bar{B}$ and denote by n_i the jump times and ν_h the stopping time representing that $\xi_n^h \in X^h \backslash B^h$. We consider a terminal state Δ and we extend the state space of the MC to $\tilde{X}^h = \bar{X}^h \cup \{\Delta\}$. The transition probabilities are defined so that $\tilde{p}^h(x, \Delta) = 1$ if $x \in X^h \setminus B^h$, $\tilde{p}^h(\Delta, \Delta) = 1$, and $\tilde{p}^h(x, x') = p^h(x, x')$ otherwise. This means that when the state reaches the unsafe set, it transitions to Δ and stays there for ever. Consider the function $\tilde{c} : \tilde{X}^h \to \mathbb{R}_+$ with $\tilde{c}(\Delta) = 1$ and $\tilde{c}(x) = 0$ for every x and the value function

$$\tilde{V}^h(x) = E_x[\sum_{n=0}^{\infty} \tilde{c}(\xi_n)]. \tag{4}$$

Clearly, this sum is well-defined, bounded. By applying the results in [21], we can show that the function $\tilde{V}^h$ can be computed using value iteration assumming appropriate initial conditions. and that the solution based on the discrete approximations converges to the one for the original stochastic hybrid system as the discretization becomes finer.

Proposition 2. *(1) Let $\tilde{V}_0^h(x) = 0$ for every x, then the iteration*

$$\tilde{V}_{n+1}^h(x) = \left[\sum_{x'} \tilde{p}^h(x, x')\tilde{V}_n^h(x')\right] \tag{5}$$

converges pointwise and monotonically to $\tilde{V}^h = V^h$.
(2) Consider the value function $V(x)$ *for the SHS, then* $\lim_{y \to x, h \to 0} \tilde{V}^h(y) = V(x)$ *uniformly in* $\bar{B}$.

Analysis of the computational complexity of value iteration algorithms is based on the contraction properties of the iteration operator. Although, the iteration operator used for verification of SHS corresponds to an undiscounted criterion, we have shown that the iteration operator restricted to an appropriate set is a contraction mapping with respect to some weighted infinity norm. Based on the contraction property, we can conclude that the iteration defined by equation (5) converges to the desired value function in a number of steps that is polynomial in the number of states N of the discrete approximating process $\{s_n^h, n = 1, \ldots, N\}$ and the number of bits used to represent the parameters of the process. Details can be found in [22].

5 Experimental Results

In this section we analyze the safety probability for the SHS sugar cataract model presented in Section 3. The chemical concentration ranges used are presented in Table 2, and the resolution of each range is presented in Table 4. We chose the resolution parameters to be similar to the resolution that measurement equipment can achieve in actual experiments. For example, the concentration of sorbitol can be experimentally measured with sub microMolar resolution.

Reactant	Resolution Step (μM)
$NADH$	1.0
$E - NADH$	1.0
NAD^+	0.5
$E - NAD^+$	0.5
sorbitol dehydrogenase (E)	0.008
fructose (F)	100
sorbitol (S)	100

Table 4. Chemical species MDP resolution for experiments

The resolution parameters for the sugar cataract system result in an MDP with approximately two billion states. Storing the values at each state alone requires several gigabytes of memory, so we developed a parallel value iteration implementation to improve the performance of the algorithm. The value iteration algorithm is still guaranteed to converge in a parallel implementation as long as updated values are used periodically [6]. Parallel dynamic programming algorithms are well-defined and easy to implement [6]. Our MDP has a regular structure which improves the efficiency of the value iteration algorithm and allows us to implement a fairly straitforward partitioning technique for the parallel implementation.

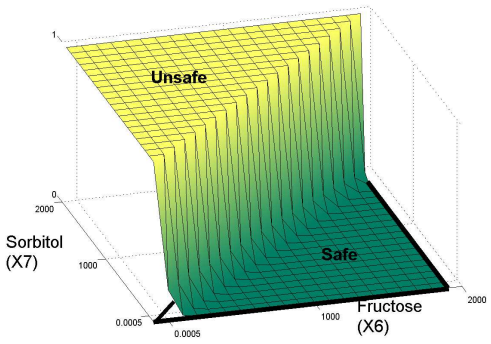

Fig. 1. Projection of the value function

To partition the problem for multiple processors we select five of the seven dimensions of the MDP to divide in half. Each processor only analyzes half of the total range for each of five divided ranges and the entire range for the other two dimensions. The two range divisions in five dimensions create $2^5 = 32$ range combinations that must be considered. The processors are each specifically assigned a combination of the ranges to ensure that the entire range for each dimension is computed, and all range values are arranged to minimize communication. Processors with neighboring range values regularly update their neighbors to ensure the value iteration converges.

To visualize our results we can plot projections of the data for different concentrations of the chemicals involved. Specifically, these projections show the safety probability for entire range of sorbitol and fructose levels for certain values of the five other variables. Multiple selections of the five other variables are chosen to show a more comprehensive view of the data.

Figure 1 shows a projection of the value function along the safety boundary where $x_1 = 1.0$, $x_2 = 20.0$, $x_3 = 1.0$, $x_4 = 0.005$, and $x_5 = 0.1$. Near the boundary of the safe and unsafe regions, the value function varies significantly depending on the projection variables chosen. This implies that certain chemical concentrations are more prone to developing cataracts than others. This data could possibly be used to help better predict sugar cataracts by demonstrating where the safest and most unsafe concentrations exist. It could also give guidance for choosing the most effective or economical treatment to avoid the cataract development.

The Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University provides the parallel computing resources for our experiments (www.accre.vanderbilt.edu). The computers form a cluster of 348 JS20 IBM PowerPC nodes running at 2.2 GHz with 1.4 Gigabytes of RAM per machine. We use C++ as the implementation language because ACCRE supports Message Passing Interface (MPI) compilers for C++. We use the MPI standard for communication between processors because it provides an efficient protocol for message passing middleware for distributed memory parallel computers. The

sugar cataract experiment took approximately 20 hours on the 32 processors. Currently, the bottlenecks of this approach are the memory size and speed.

6 Conclusions

Biochemical system modeling and analysis are important but challenging tasks which hold promise to unlock secrets of complicated biochemical systems. SHS are an ideal modeling paradigm for biochemical systems because they incorporate probabilistic dynamics into hybrid systems to capture the inherent stochastic nature of the biochemical systems. The sugar cataract development problem is excellent example of a system that is modeled effectively using the presented modeling methods. Our dynamic programming analysis technique provides verification results for realistic systems using parallel computing techniques to lessen the effect of the curse of dimensionality.

Acknowledgements

We would like to thank Howard Salis and Yiannis Kaznessis from the University of Minnesota for their help with this project.

References

1. R. Alur, C. Belta, F. Ivanicic, V. Kumar, M. Mintz, G. Pappas, H. Rubin, J. Schug, Hybrid Modeling and Simulation of Biomolecular Networks, *Hybrid Systems: Computation and Control*, LNCS 2034, pp. 19-33, 2001.
2. S. Amin, A. Abate, M. Prandini, J. Lygeros, S. Sastry, Reachability Analysis for Controlled Discrete Time Stochastic Hybrid Systems, *Hybrid Systems: Computation and Control 2006*, LNCS 3927, pp. 49-63, 2006.
3. R. Barbuti, A. Maggiolo-Schettini, P. Milazzo, A. Triona, An Alternative to Gillespie's Algorithm for Simulating Chemical Reactions, *Computational Methods in Systems Biology 2005*, Edinburgh, 2005.
4. R. Barbuti, S. Cataudella, A. Maggiolo-Schettini, P. Milazzo, A. Triona, A Probabilistic Model For Molecular Systems, *Fundamenta Informaticae XX, IOS Press*, 1-15, 2005.
5. M. Bernadskiy, R. Sharykin, and R. Alur, Structured Modeling of Concurrent Stochastic Hybrid Systems, *FORMATS'04*, LNCS 3253, pp. 309-324, 2004.
6. D. Bertsekas, J. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods, Prentice-Hall, 1989.
7. M. Bujorianu, J. Lygeros, Theoretical Foundations of General Stochastic Hybrid Systems: Modeling and Optimal Control, *In Proc. of 43rd IEEE Conf. on Decision and Control*, 2004.
8. M. Bujorianu, J. Lygeros, Reachability Questions in Piecewise Deterministic Markov Processes, *Hybrid Systems: Computation and Control 2003*, LNCS, vol. 2623, pp. 126-140, 2003.
9. Y. Cao, H. Li, L. Petzold, Efficient formulation of the stochastic simulation algorithm for chemically reacting systems, *J. Chem. Phys.*, 121, 4059, 2004.
10. M. Davis, Markov Models and Optimization, Chapman and Hall, 1993.
11. M. Elowitz, A. Levine, E Siggia, P. Swain, Stochastic Gene Expression in a Single Cell, *Science 297*, 1183, 2002.

12. R. Ghosh, C. Tomlin, Symbolic Reachable Set Computation of Piecewise Affine Hybrid Automata and its Application to Biological Modeling: Delta-Notch Protein Signalling, *Systems Biology*, 1:170-183, 2004.
13. D. Gillespie, A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *J Comp. Phys.*, 22:403-434, 1976.
14. E. Haseltine, J. Rawlings, Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics, *J Chem. Phys.*, 117, 6959, 2002.
15. J. Hespanha, Stochastic Hybrid Systems: Application to Communication Networks, *Hybrid Systems: Computation and Control 2004*, LNCS, vol. 2993, pp. 387-401, 2004.
16. J. Hespanha, A. Singh. Stochastic Models for Chemically Reacting Systems Using Polynomial Stochastic Hybrid Systems, *Int. J. on Robust Control, Special Issue on Control at Small Scales*, 15:669689, 2005.
17. J. Hu, M. Prandini, and S. Sastry, Probabilistic Safety Analysis in Three Dimensional Aircraft Flight, *Proc. of 42nd IEEE Conf. on Decision and Control*, pp. 5335-5340, 2003.
18. J. Hu, W. Wu, and S. Sastry, Modeling Subtilin Production in Bacillus subtilis Using Stochastic Hybrid Systems, *Hybrid Systems: Computation and Control 2004*, LNCS, vol. 2993, pp. 417-431, 2004.
19. C. Morton-Firth, Stochastic Simulation of Cell Signaling Pathways. PhD Thesis. University of Cambridge, 1998.
20. X. Koutsoukos, Optimal Control of Stochastic Hybrid Systems Based on Locally Consistent Markov Decision Processes, *Int. J. of Hybrid Systems*, 4, 301-318, 2004.
21. X. Koutsoukos and D. Riley, Computational Methods for Reachability Analysis of Stochastic Hybrid Systems, *Hybrid Systems: Computation and Control 2006* LNCS 3927, pp. 377-391, 2006.
22. X. Koutsoukos and D. Riley, Computational Methods for Verification of Stochastic Hybrid Systems, submitted.
23. M. Kwiatkowska, G. Norman, D. Parker, O. Tymchyshyn, J. Heath, E. Gaffney, Simulation and Verification for Computational Modeling of Signaling Pathways. *Proceedings of the Winter Simulation Conference*, to appear Dec 2006.
24. H. Kushner, P. Dupuis, Numerical Methods for Stochastic Control Problems in Continuous Time, Springer, 2001.
25. J. Lygeros, On reachability and minimum cost optimal control, *Automatica*, 40(6), 917-927, 2004.
26. I. Marini, L. Bucchioni, P. Borella, A. Del Corso, U. Mura, Sorbitol Dehydrogenase from Bovine Lens: Purification and Properties, *Archives of Biochemistry and Biophysics*, 340:383-391, 1997.
27. I. Mitchell, A. Bayen, and C. Tomlin, A Time-Dependent Hamilton-Jacobi Formulation of Reachable Sets for Continuous Dynamic Games, *IEEE Trans. on Automatic Control*, 50(7), 947-957, 2005.
28. I. Mitchell, J. Templeton, A Toolbox of Hamilton-Jacobi Solvers for Analysis of Nondeterministic Continuous and Hybrid Systems, *Hybrid Systems: Computation and Control 2005*, LNCS, vol. 3414, pp. 480-494, 2005.
29. M. Prandini, O. Watkins, Probabilistic Aircraft Conflict Detection, *HYBRIDGE, IST-2001-32460*, 2005.
30. M. Puterman, Markov Decision Processes-Discrete Stochastic Dynamic Programming, Wiley: Hoboken, New Jersey, 2005.
31. H. Salis and Y. Kaznessis, Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions, *The Journal of Chemical Physics*, v122, pp. 54-103, 2005.