SAND2007-6817C

Addressing biological circuit simulation accuracy: Reachability for parameter identification and initial conditions

Meeko Oishi[†], Elebeoba May[‡]
[†]Electrical and Computer Engineering, University of British Columbia, Vancouver, BC
moishi@ece.ubc.ca
[‡]Computational Biology, Sandia National Laboratories, Albuquerque, NM
eemay@sandia.gov

***Abstract*— Accurate simulation of biological networks is difficult not only due to the computational cost associated with large-scale systems simulation, but also due to the inherent limitations of mathematical models. We address two components to improve biological circuit simulation accuracy: 1) feasible initial conditions, and 2) identification of critical yet unknown model parameters. For those parameters that may not be available from experimental data, we incorporate reachability analysis to enhance our optimization/simulation framework and estimate those parameters that are capable of creating behaviors consistent with known experimental data. We apply these techniques to a biological circuit model of tryptophan biosynthesis in *E. coli*, and quantify the improvement in simulation accuracy when reachability analysis is used.**

Keywords: **hybrid systems, reachability, biological circuits, tryptophan, Escherichia coli, parameter identification.**

I. Introduction

Accurately modeling and simulating biological networks is a daunting problem, due to the complex interaction between large numbers of interacting pathways, feedback inherent to the system, and the stochastic nature of biological processes. However, recent techniques have been developed to model and simulate large-scale biological networks using analogies to electrical circuits [1]. Exploiting the similarities between biological networks and electrical circuit networks, an efficient parallel circuit simulator, *Xyce* [2], is used to simulate large-scale circuit equivalents of biological processes. The result is computational tool to model and simulate multivariate, multiscale, hybrid biological networks.

However, as with any complex simulation, the results of the computation are highly dependent not only on the numerical accuracy of the simulation technique, but also on the particular values of model parameters as well as the simulation's initial conditions. Biological circuits are inherently hybrid, with both discrete and continuous components. Hybrid systems are notorious for their non-intuitive behavior, and potentially high sensitivity to variations in initial conditions. Blindly choosing unknown initial conditions and parameters may make it impossible to simulate the desired behavior – for example, it may not be possible for a model to reach a desired equilibrium from a given initial condition, even though the mechanics of the model are correct. Additionally, biological parameters that are not known experimentally (incomplete or noisy data), may have an unintentionally large effect on biological simulation accuracy.

We present in this paper a method to address these two issues: 1) unknown initial conditions for biological circuit simulations, and 2) inaccurate or unknown model parameters, by incorporating computational techniques for reachability analysis. To solve the first issue we wish to find the set of *feasible* initial conditions – that is, those initial conditions from which it is possible to achieve the desired system behavior (e.g., an experimentally-determined steady-state protein concentration). To solve the second issue we wish to find combinations of parameter values (e.g., unknown or erroneous metabolic rate constants) that will result in the desired system behavior. These two problems are typically coupled, however, due to special structure in the metabolic equations we use, these two problems can be solved independently of each other.

A variety of computational approaches, based on optimization [3], [4], evolutionary algorithms [5], and others, have been used to estimate biological parameters. However, *reachability*, or the ability of a system to 'reach' a desired set of states, provides a highly accurate, though often computationally intensive, description of the system's behavior [6]. As opposed to simulation, in which trajectories are evolved from a finite set of initial conditions, reachability analysis provides a mathematical guarantee over an infinite number of initial conditions. Reachability tools have been primarily used in verification problems in engineering applications, however more recently these tools have been adapted for use in parameter identification [7] and model validation [8] of biological processes.

This paper provides a technique for solving 1) the initial conditions problem, and 2) the parameter problem. We implement these techniques on a simple example: the tryptophan biosynthesis network in *Escherichia coli* K-12. In Section II the circuit modeling technique is briefly de-

Thanks to Richard L. Schiek and Brian M. Adams for their assistance with Xyce and DAKOTA toolkit. The authors contributed equally to this work. This work was supported by an NSERC Discovery Grant, a Harry S. Truman Postdoctoral Fellowship in National Security Science and Engineering at Sandia National Laboratories, and NIH/NHLBI grant 5K25HL75105-3. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company for the US Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

scribed. Section III addresses the problem of feasible initial conditions, and Section IV describes our use of reachability techniques to compute critical unknown parameters. Finally, we present our results with this method for the *E. coli* tryptophan biosynthesis system, and conclude with directions for future work.

II. Biological circuit modeling

In order to take advantage of the *Xyce* simulation framework, we create circuit abstractions of biological elements and construct netlist files that are executed using *Xyce* [9]. Similar to the abstractions used in flux balance analysis, the flow of metabolic and genetic substrates are synonymous to the flow of current through an electrical circuit [10], [11]. Metabolic reactions are simulated using analog sub-circuits, where metabolite accumulation and degradation are modeled using capacitors and resistors, respectively.

A. Tryptophan Biosynthesis

A major challenge in the development of accurate biological network simulations is the availability of accurate rate data for metabolic reactions. We incorporate data from the BRENDA database (*www.brenda.uni-koeln.de*), which contains empirically determined reaction rates for organisms including *E. coli*. The stoichiometric reactions for tryptophan biosynthesis [12], [13] in *E. coli* have enzyme EC numbers

$$\begin{array}{ll} \text{(EC 4.1.3.27)} & Chor + Gln \rightarrow Glu + Pyr + An \\ \text{(EC 2.4.2.18)} & An + Prpp \rightarrow Ppi + Npran \\ \text{(EC 5.3.1.24)} & Npran \rightarrow Cpad5p \\ \text{(EC 4.1.1.48)} & Cpad5p \rightarrow Co2 + Igp \\ \text{(EC 4.2.1.20)} & Igp + Ser \rightarrow T3p1 + Trp \end{array} \tag{1}$$

with metabolites abbreviated as Chor: Chorismate, Gln: Glutamine, Glu: Glutamate, Pyr: Pyruvate, An: Antranilate, Prpp: Phosphoribosyl pyrophosphate, Ppi: Pyrophosphate, Npran: N-(5'-phosphoribosyl)-anthranilate, Cpad5p: 1-(O-Carboxyphenylamino)-1'-deoxyribulose-5'phosphate, Co2: Carbon dioxide, Igp: Indole glycerol phosphate, Ser: Serine, T3p1:Glyceraldehyde 3-phosphate, Trp: Tryptophan.

We use the Michaelis-Menten equation to describe the reaction kinetics with the Michaelis-Menten constant, K_M, and the turnover rate, K_{cat} [14]. The Michaelis-Menten equation for an enzyme catalyzed reaction relates the rate at which the product P is made to rate at which its precursor substrate S is modified by the enzyme E. The equation for a single substrate/single product reactions $E + \text{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{ES} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \text{E} + P$ assumes that the breakdown of the enzyme-substrate complex to product, $[ES] \rightarrow [P]$, is the rate limiting step. Equating K_{cat} to k_2, the reaction rate is

$$V = \frac{dP}{dt} = \frac{K_{\text{cat}}^S [S]}{K_M^S + [S]} \tag{2}$$

where $K_M = \frac{k_{-1}+k_2}{k_1}$. We extend this analysis to multiple substrate/multiple product systems by maintaining the steady state and constant total enzyme concentration assumptions, and further assuming that the breakdown of the enzyme-substrates complex is the limiting step. These assumptions cause the rate equation to depend mainly on the number of inputs. For the double substrate reactions in (1), a second generic rate equation is derived using the King-Altman method (*www.biokin.com/king-altman/theory.html*):

$$V = \frac{dP}{dt} = \frac{K_{\text{cat}}^{S_1 S_2} [S_1][S_2]}{[S_1][S_2] + [S_1] K_M^{S_2} + K_M^{S_1} K_M^{S_2}} \tag{3}$$

These equations explicitly incorporate kinetic parameters reported in *BRENDA* for *E. coli*, queried by the enzyme EC numbers in (1). When possible, we selected rates curated from the same reference document and those derived using a wildtype, non-mutant strain. With data for for all but five of the thirteen rate parameters, the known Michaelis-Menten rates are $K_M^{\text{Chor}} = 1.2e^{-3}$, $K_M^{\text{Gln}} = 3.16e^{-1}$, $K_M^{\text{Prpp}} = 1.3e^{-2}$, $K_M^{\text{Npran}} = 4.9e^{-3}$, $K_M^{\text{Cpad5p}} = 3.4e^{-4}$, and the known turnover rates are $K_{\text{cat}}^{\text{AnPrpp}} = 4.4$, $K_{\text{cat}}^{\text{Npran}} = 50.0$, $K_{\text{cat}}^{\text{Cpad5p}} = 2.2$. The unknown rates $K_M^{\text{An}}, K_M^{\text{Igp}}, K_M^{\text{Ser}}, K_{\text{cat}}^{\text{ChorGln}}, K_{\text{cat}}^{\text{IgpSer}}$ are addressed in Sections IV and V.

B. Dynamical model

We incorporate the metabolic rates into the Michaelis-Menten equations (2), (3), resulting in an ODE with state $x \in \mathcal{R}^9$, representing the concentrations of eight proteins involved in tryptophan biosynthesis.

$$\begin{bmatrix} \dot{x}_{\text{Chor}} \\ \dot{x}_{\text{Gln}} \\ \dot{x}_{\text{An}} \\ \dot{x}_{\text{Prpp}} \\ \dot{x}_{\text{Npran}} \\ \dot{x}_{\text{Cpad5p}} \\ \dot{x}_{\text{Igp}} \\ \dot{x}_{\text{Ser}} \\ \dot{x}_{\text{Trp}} \end{bmatrix} = \begin{bmatrix} -f_A(x_{\text{Chor}}, x_{\text{Gln}}) \\ -f_A(x_{\text{Chor}}, x_{\text{Gln}}) \\ f_A(x_{\text{Chor}}, x_{\text{Gln}}) - f_B(x_{\text{An}}, x_{\text{Prpp}}) \\ -f_B(x_{\text{An}}, x_{\text{Prpp}}) \\ f_B(x_{\text{An}}, x_{\text{Prpp}}) - f_C(x_{\text{Npran}}) \\ f_C(x_{\text{Npran}}) - f_D(x_{\text{Cpad5p}}) \\ f_D(x_{\text{Cpad5p}}) - f_E(x_{\text{Igp}}, x_{\text{Ser}}) \\ -f_E(x_{\text{Igp}}, x_{\text{Ser}}) \\ f_E(x_{\text{Igp}}, x_{\text{Ser}}) \end{bmatrix} \tag{4}$$

Note that $f_A(x) = \frac{K_{\text{cat}}^{\text{ChorGln}} x_{\text{Chor}} x_{\text{Gln}}}{x_{\text{Chor}} x_{\text{Gln}} + K_M^{Gln} x_{\text{Chor}} + K_M^{\text{Gln}} K_M^{\text{Chor}}}$, $f_B(x) = \frac{K_{\text{cat}}^{\text{AnPrpp}} x_{\text{An}} x_{\text{Prpp}}}{x_{\text{An}} x_{\text{Prpp}} + K_M^{\text{Prpp}} x_{\text{An}} + K_M^{\text{An}} K_M^{\text{Prpp}}}$, $f_C(x) = \frac{K_{\text{cat}}^{\text{Npran}} x_{\text{Npran}}}{x_{\text{Npran}} + K_M^{\text{Npran}}}$, $f_D(x) = \frac{K_{\text{cat}}^{\text{Cpad5p}} x_{\text{Cpad5p}}}{x_{\text{Cpad5p}} + K_M^{\text{Cpad5p}}}$, and $f_E(x) = \frac{K_{\text{cat}}^{\text{IgpSer}} x_{\text{Igp}} x_{\text{Ser}}}{x_{\text{Igp}} x_{\text{Ser}} + K_M^{\text{Ser}} x_{\text{Igp}} + K_M^{\text{Igp}} K_M^{\text{Ser}}}$.

III. Feasible Initial Conditions

These dynamics have stable equilibria at x^* which satisfies

$$0 = x^*_{\text{Chor}} x^*_{\text{Gln}} = x^*_{\text{An}} x^*_{\text{Prpp}} = x^*_{\text{Npran}} = x^*_{\text{Cpad5p}} = x^*_{\text{Igp}} x^*_{\text{Ser}} \tag{5}$$

The conservation of substrate appears in (4) as the summation of various combinations of rates to 0:

$$0 = A\dot{x}, \quad A = \begin{bmatrix} 1 & -1 & & & & & & & & \\ & 1 & 1 & -1 & & & & & & \\ & & & & 1 & 1 & 1 & 1 & -1 & \\ & & & & & & & & 1 & -1 \end{bmatrix} \tag{6}$$

Integrating (6) as $t \rightarrow \infty$, we obtain restrictions on the initial condition $x(0)$

$$Ax(0) = b, \tag{7}$$

where $b = Ax^*$. For experimental data [15] with $x^*_{\mathrm{Trp}} = 395$, $b = [0, 0, 0, 395]^T$. Any initial conditions which satisfy (7) for this value of b will reach the desired concentration of tryptophan in steady-state. Thus we have determined the set of initial conditions will result in a desired steady-state behavior (7). This analysis guides the choice of initial voltage value across the capacitors connected to each substrate node.

The dynamical system (4) has the interesting property that its steady-state behavior is dependent on the initial conditions, and its transient behavior is dependent on the rate parameters. This decoupling allows us to *independently* investigate 1) feasibility of initial conditions, and 2) identification of unknown parameters.

IV. Parameter identification

Reachability analysis provides a synthesis of the largest set of parameters which will match known experimental data. As opposed to optimization, which may become "stuck" in local minima, reachability techniques fully explore the state-space and therefore provide more complete information about the system behavior. We use these techniques only for those parameters which are critical for matching known data.

Computing the reachable set involves enumerating all the states which have a path to a target set. We draw on level set methods here because of their sub-grid accuracy and success in previous aircraft applications [16], [6], however other techniques can be used [17], [18]. Typically used for safety verification, we use reachability in a different way here: by extending the state-space with unknown parameters of interest, we can synthesize the largest set of parameter values which lead to a desired target in the state-space [19].

This is accomplished by a computation in backwards time: starting with the desired target, which is encoded implicitly as a level set function, the boundary of the target set is propagated backwards in time according to the system dynamics. To calculate the reachable set $\mathcal{W}(t)$, first define a continuous function $J_0 : \mathcal{X} \to \mathbb{R}$ which encodes the target

$$\mathcal{W}_0 = \{x \in \mathcal{X} \mid J_0(x) \leq 0\}. \tag{8}$$

Finding the backwards reachable set $\mathcal{W}(t)$ requires solving the terminal value time-dependent modified Hamilton-Jacobi (HJ) partial differential equation (PDE) [19]

$$\begin{aligned} \frac{\partial J(x,t)}{\partial t} + \min\left[0, H\left(x, \frac{\partial J(x,t)}{\partial x}\right)\right] = 0 \qquad & \text{for } t < 0; \\ J(x,0) = J_0(x) \qquad & \text{for } t = 0; \end{aligned} \tag{9}$$

with $H(x, \frac{\partial J(x,t)}{\partial x}) = \frac{\partial J(x,t)}{\partial x}^T f(x)$. As shown in [19], we obtain an implicit representation of the reachable set $\mathcal{W}(t) = \{x \in \mathcal{X} | J(x,t) \leq 0\}$.

In the parameter problem, the reachable set $\mathcal{W}(t)$ represents those combinations of state and parameters which reach the desired protein concentrations in the time t. Consider a simplified model of tryptophan biosynthesis,

$$\begin{bmatrix} \dot{x}_{\mathrm{Ser}} \\ \dot{x}_{\mathrm{Trp}} \end{bmatrix} = \begin{bmatrix} -K^{\mathrm{IgpSer}}_{\mathrm{cat}} \frac{x_{\mathrm{Ser}}}{x_{\mathrm{Ser}} + K^{\mathrm{Ser}}_M} \\ K^{\mathrm{IgpSer}}_{\mathrm{cat}} \frac{x_{\mathrm{Ser}}}{x_{\mathrm{Ser}} + K^{\mathrm{Ser}}_M} \end{bmatrix} \tag{10}$$

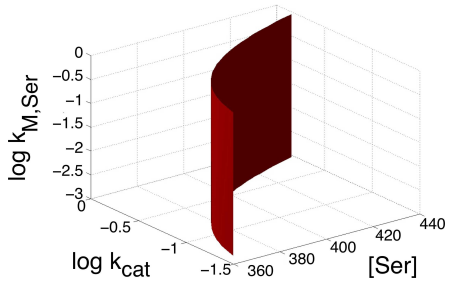

Fig. 1. Combinations of x_{Ser}, $K^{\mathrm{IgpSer}}_{\mathrm{cat}}$, and K^{Ser}_M that achieve the target $x_{\mathrm{Trp}}(t_f) = 35$, $t_f = 120$ s.

obtained by assuming that $K^{\mathrm{Igp}}_M K^{\mathrm{Ser}}_M \ll x^{\mathrm{Igp}}_M x^{\mathrm{Ser}}_M + K^{\mathrm{Ser}}_M x_{\mathrm{Igp}}$. We extend (10) with $\dot{K}^{\mathrm{Ser}}_M = \dot{K}^{\mathrm{IgpSer}}_{\mathrm{cat}} = 0$ to compute the largest set of $K^{\mathrm{Ser}}_M, K^{\mathrm{IgpSer}}_{\mathrm{cat}}$ for which the system will achieve trajectories in x_{Trp} which match experimental data. Of the five unknown paramters, the two parameters that appear in (10) are considered 'critical' for matching the experimental data; the remaining three are not critical.

Since $x_{\mathrm{Trp}}(t) = -x_{\mathrm{Ser}}(t) + x_{\mathrm{Ser}}(0) + x_{\mathrm{Trp}}(0)$, constraints on x_{Trp} can be written in terms of x_{Ser}. Therefore $x_{\mathrm{Trp}}(t_f) = 35$ at $t_f = 120$ seconds, with $x(0) = [0, 395]$ is equivalent to a target set $\mathcal{W}_0 = \{x \mid x_{\mathrm{Ser}} \leq 360\}$.

The result is shown in Figure 1. All points on the red surface are those combinations of x_{Ser}, $\log K^{\mathrm{IgpSer}}_{\mathrm{cat}}$, $\log K^{\mathrm{Ser}}_M$ which achieve $x_{\mathrm{Ser}}(t_f) = 360$. Surprisingly, K^{Ser}_M does not seem to significantly influence the trajectories, while $K^{\mathrm{IgpSer}}_{\mathrm{cat}}$ does. Examining Figure 1, we see that trajectories which start from $x_{\mathrm{Ser}}(0) = 395$ and end at the target in the allotted time require a value of $K^{\mathrm{ChorGln}}_{\mathrm{cat}} \approx 0.3$.

V. An optimization framework for parameter estimation and refinement

As the number of metabolites and reactions in the biological system increases, so does the number of kinetic parameters needed for reliable system simulation. To efficiently estimate large numbers of unknown or unreliable parameters we couple *Xyce* with the *DAKOTA* (Design Analysis Kit for Optimization and Terascale Applications) toolkit, which provides a parallel platform for optimization [20]. We create a loosely-coupled interface between DAKOTA and the *Xyce* simulator through a parameter file generated by DAKOTA based on response data derived from simulation results. We use this coupling to refine known reaction rates and determine the optimal rate constants for the five unknown rate constants in our *in silico* model of tryptophan biosynthesis.

We compare empirical data from Bliss 1979 [15] to 3600 seconds of *Xyce* simulation data. For known reaction rates we initialize the rate parameter to its reported value(s) and allow DAKOTA to search within a bounded region of the same order of magnitude that contains the reported value(s). For unknown rates, we initialize the K_M to 0.05 mM and the K_{cat} to 50 s^{-1}, and bound DAKOTA's search within $[1e^{-4}, 1e^{0}]$ mM and $[1e^{0}, 1e^{2}]$ s^{-1}, respectively. DAKOTA minimizes a quadratic objective function $J(x) = \sum_{i=1}^{12} [T_i(x)]^2$, where $T_i(x)$ is the difference between the empirically determined

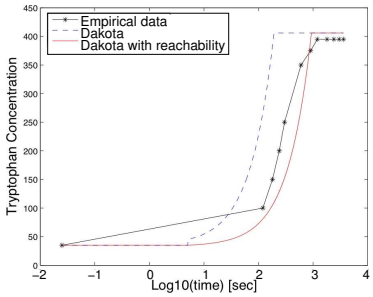

Fig. 2. Tryptophan concentrations over log-time for empirical data (points), optimization without reachability (dashed), and optimization with reachability (solid).

concentration (equated to measured fluorescence intensity) and the *Xyce* predicted average tryptophan concentration at time i. DAKOTA instantiates multiple runs of *Xyce* until it finds a parameter set that best minimizes the objective function or it reaches a predefined stopping criteria (e.g. maximum number of iterations).

To evaluate the efficacy and impact of reachability on the parameter estimation problem, we implement the coupled optimization/simulation framework under two assumptions: 1) Unavailability of reachability information; 2) Availability of reachability information (where $K_{\text{cat}}^{\text{IgpSer}} = 0.3\text{s}^{-1}$ initially, bounded by $[0.2, 0.4]$ s^{-1}). Figure 2 shows the evolution of tryptophan concentration over time for the empirical data and *Xyce* simulations using the optimal parameters returned by DAKOTA. As Figure 2 shows, the incorporation of the reachability results improved simulation accuracy. The least-squares difference between the empirical and *in silico* models was reduced by a factor of 6.5 when the reachability analysis was incorporated. The optimization with reachability completed 260 runs, with optimal values found on run 243; while the optimization without reachability information completed 61 runs, with optimal values found on run 45. The limited number of runs for the non-reachability case may be due to DAKOTA becoming trapped in a local minima, whereas the implementation with reachability may have avoided this by identifying a plausible value for the critical system parameter. The reachability analysis took 10:17:37.6 hours to run 32262 steps on an Intel-based Powerbook.

VI. Conclusion

Parameter identification is a challenge in the pursuit to produce accurate and relevant *in silico* models and simulations of biological systems. The number of kinetic parameters increases exponentially as we attempt to produce reliable simulations of whole cells and multicellular systems. Coupling the massively parallel *Xyce* with DAKOTA and available empirical data provides an initial framework for elucidating unknown reaction rates. As demonstrated in this work, the incorporation of reachability analysis increases the effectiveness of the coupled DAKOTA/*Xyce* framework and results in more reliable models for systems biology applications. Future work involves extension of the approaches described to refine whole-cell and multicellular network models for the simulation of central metabolism in *E. coli* K-12. Challenges include solving the initial condition problem for a significantly larger set of metabolites, extending the reachability methodologies to larger systems, automation, and parallel application of reachability techniques. Successful treatment of biological networks as engineering systems in modeling and simulation can reduce research and development cost, enable pharmaceutical target discovery, and provide insight into the regulation of numerous cellular processes.

References

[1] E. May and R. Schiek, "Xyce parallel electronic simulator: Biological pathway modeling and simulation," SAND Report SAND2006-1993P, Sandia National Laboratories, Tech. Rep., 2006.

[2] S. D. Wix, E. R. Keiter, S. A. Hutchinson, R. J. Hoekstra, E. L. Rankin, R. P. Pawlowski, D. A. Fixel, T. V. Russo, and L. J. Waters, "Xyce parallel electronic simulator: Reference guide," Sandia Nat'l Laboratories, Albq., NM, SAND2004-2896, 2004, `www.cs.sandia.gov/xyce`.

[3] P. Henning, R. Moffitt, J. Allegood, E. Wang, A. Merrill, and M. Wang, "Computationally predicting rate constants in pathway models," in *Proc. IEEE Int'l Eng. Med. & Bio. Conf.*, Sept. 2005, pp. 5093–5096.

[4] R. Raffard, K. Amonlirdviman, J. Axelrod, and C. Tomlin, "Automatic parameter identification via the adjoint method, with application to understanding planar cell polarity," in *Proc. IEEE Conf. on Decision and Control*, San Diego, CA, Dec. 2006, pp. 13–17.

[5] P. Gennemark and D. Wedelin, "Efficient algorithms for ordinary differential equation model identification of biological systems," *IET Systems Biology*, vol. 1, no. 2, pp. 120–129, 2007.

[6] I. Mitchell, *A Toolbox of Level Set Methods*, Department of Computer Science, University of British Columbia, June 2004, www.cs.ubc.ca/∼mitchell/ToolboxLS.

[7] M. Kloetzer and C. Belta, "Reachability analysis of multi-afine systems," in *Hybrid Systems: Computation and Control*, ser. LNCS 3927, J. Hespanha and A. Tiwari, Eds. Springer, March 2006, pp. 348–362.

[8] H. El-Samad, S. Prajna, A. Papachristodoulou, J. Doyle, and M. Khammash, "Advanced methods and algorithms for biological networks analysis," *Proceedings of the IEEE*, vol. 94, no. 4, pp. 832–853, 2006.

[9] E. May and R. Schiek, "Simulating regulatory networks using $Xyce^{\text{TM}}$," in *Int'l Conf. on Systems Biology*, 2003.

[10] C. H. Schilling, J. S. Edwards, D. Letscher, and B. Palsson, "Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems," *Biotechnology and Bioengineering*, vol. 71, no. 4, pp. 286–306, 2001.

[11] M. W. Covert, C. H. Schilling, and B. Palssson, "Regulation of Gene Expression in Flux Balance Models of Metabolism," *Journal of Theoretical Biology*, vol. 213, pp. 73–88, 2001.

[12] SRI, "BioCyc database collection," `www.biocyc.org`.

[13] B. Palsson, "*E. Coli* reconstruction files," `gcrg.ucsd.edu`.

[14] D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry, 4th Edition*. New York: W. H. Freeman and Company, 2005.

[15] R. D. Bliss, "A specific method for determination of free tryptophan and endogenous tryptophan in *Escherichia coli*," *Analytical Biochemistry*, vol. 93, pp. 390–398, 1979.

[16] M. Oishi, I. Mitchell, A. Bayen, C. Tomlin, and A. Degani, "Hybrid verification of an interface for an automatic landing," in *Proc. IEEE Conf. Decision & Control*, Las Vegas, NV, Dec. 2002, pp. 1607–1613.

[17] A. Chutinan and B. Krogh, "Computational techniques for hybrid system verification," *IEEE Trans. Automatic Control*, vol. 48, no. 4, pp. 64–75, Jan. 2003.

[18] E. Asarin, T. Dang, and A. Girard, "Reachability analysis of nonlinear systems using conservative approximation," in *Hybrid Systems: Computation and Control*, ser. LNCS 2623, O. Maler and A. Pneuli, Eds. Springer Verlag, March 2003, pp. 20–35.

[19] C. Tomlin, I. Mitchell, A. Bayen, and M. Oishi, "Computational techniques for the verification of hybrid systems," *Proceedings of the IEEE*, vol. 91, no. 7, pp. 986–1001, 2003.

[20] M. Eldred, A. Giunta, S. Brown, B. Adams, D. Dunlavy, J. Eddy, D. Gay, J. Griffin, W. Hart, P. Hough, T. Kolda, M. Martinez-Canales, L. Swiler, J. Watson, and P. Williams, "DAKOTA, A multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis," Sandia Nat'l Laboratories, Albq., NM, SAND2006-4055, 2006.