Synthesis of Optimal Control of Complex Biological Pathways Enabled by Global Sensitivity Analysis

Amin Salighehdar[1] and Narayan Ganesan[2]

Abstract— **The past few years have witnessed growth in the number of system biological models corresponding to several different pathways and have shed light on biological processes governing vital functions such as signal transduction, cell-proliferation and cell-apoptosis. Among several challenges in modeling and verification, significant efforts have been made in system identification including model parameter estimation and key component identification. In a practical or experimental setting, the effects of various control strategies, are usually associated with costs including enzyme consumption or the desired levels of output transcription. In this work we address the problem of designing an optimal control law via a combination of computational tools such as sensitivity analysis, Pontryagin maximum principle and steepest descent technique. We find that, in the presence of limitation of enzyme concentration level control, optimal control law can be developed only by choosing enzyme with dominant effect on dynamical behavior of biological pathway. In this work, we study JAK/STAT signal transduction pathway, and by simulation analysis, we investigate the effectiveness of Sobol values in developing optimal control law.**

I. INTRODUCTION

The past few years, mathematical modeling of biological pathway has become an important topic among researchers. A very useful modeling tool that captures the mechanistic and kinetic description of the interaction is Ordinary differential Equations (ODEs)[1]. Modeling complex Bio-Molecular pathways via ODEs help gain rapid insights into the dynamics of processes, nominal levels and activation times. However, the kinetic model parameters for the ODE are estimated based on the incomplete and extremely sparse experimental data which is the norm in practical biological wet-lab experiments.

As mentioned above, estimating unknown parameters always suffers uncertainties due to lack of experimental data or modeling techniques. So studying the effect of these variations on output of system can provide comprehensive insight into dynamics. To this end, sensitivity analysis can be helpful, which can be described as investigation of the model responses to perturbations of the model quantitative factors. Typically mathematical model of systems contain large number of parameters with uncertainties in their values. Sensitivity analysis methods can be categorized into two groups: local and global method [2]. In local methods, individual input is varied at a time around its nominal value, then effect of its variation on the output is analyzed. In global sensitivity analysis, the comprehensive effects of all inputs acting on the system is studied by varying them simultaneously around their nominal value. Among global sensitivity analysis techniques, variance based methods are model-free as they are not dependent on assumption about the relationship between model inputs and outputs. Variance based methods such as Fourier Amplitude Sensitivity Test(FAST), extended FAST(eFAST) and Sobol sensitivity indices are presented in [3] which are computationally expensive but are capable of determining effects between individual inputs on output, as well as total effect including all possible interactions between parameters. In [4], an application of sensitivity analysis in biological system is presented, which revealed critical reactions in signaling pathway. In [5], Sobol sensitivity analysis for complex environmental model is studied. The result of this analysis is identifying key parameters of dynamical system which play essential role in output variation. In [6], new correlation-based numerical technique to compute Sobol values is proposed, which also reduces errors associated with correlation.

In addition to parameter analysis, output regulation and control of these pathways are of importance in applications requiring optimal behavior of the pathways, as in treatment of diseases or laboratory experiments. Control and regulation of pathways were studied in [7]. Any effort involving control or regulation also involves optimizing cost of reagents, time or effort which is vital to eliciting the desired behavior from the system. A complex bio-chemical pathway typically involves several dozen input and output components, which makes it computationally intensive to search for optimality in the control space. The problem is exacerbated by the fact that the experimental biologist has the choice to include a certain component enzyme or a protein as a part of the input to the experiment, which has the ability to influence the cost function. In such a scenario choice has to be made as to which enzyme or set of enzymes or proteins serves as an optimal control set that minimizes the overall cost function. In this work, we employ global sensitivity analysis technique to determine the optimal set of enzymes that can influence the overall system behavior, which is then utilized to design optimal control law with a combination of Pontryagin's maximum principle and steepest descent techniques. The biological application driving our study is interferon-γ (IFN-γ) induced Janus Kinase-signal transducers and activators of transcription (JAK-STAT) pathway. The rest of this paper is organized as follows: In section II, JAK/STAT signaling pathway activation process is reviewed. In section III, simulation results for mathematical model, sensitivity

[1]Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA asalighe@stevens.edu

[2]Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA nganesan@stevens.edu

978-1-4244-7929-0/14/$26.00 ©2014 European Union

analysis as well as optimal control design are presented. Finally, in section IV, and V we discuss and summarize the main results of this work.

II. JAK/STAT Signaling Pathway

The Janus kinase/Signal transducer and activators of transcription (JAK/STAT) pathway transmit the signals from the extracellular media into the nucleus leading to DNA transcription. JAK/STAT pathway is the major mechanism which simulates cell proliferation, differentiation, and apoptosis. Here we illustrate the primary mechanism of activation. More detail can be found in [8]. Fig. 1, shows the dynamic scheme of JAK/STAT signaling pathway which is composed of cascade of chemical reactions [8]. At first step, JAK binds to the IFN-γ receptor (IFNR) and forms the IFNR-JAK complex (RJ in kinetic scheme). In step 2, IFN-γ binds to extracellular domain of RJ complex and forms the IFN-γ-IFNR-JAK complex (IFNRJ). In next step, dimerization of IFNRJ leads to formation of IFRJ2. The dimerization of the RJ complex leads to the phosphorylation of several tyrosine residues by JAK (Step 4), yielding IFRJ2*. In step 5, STAT1c binds to IFNRJ2* and is phosphorylated by JAK. The homodimer of phosphprylated STAT1c is formed in step 6. In next two steps (7 and 8) translocation of STAT1c dimer to nucleus is followed by the transcription of DNA. The corresponding mRNA is translated as the SOCS1 protein(step 9), which binds to the activated receptor-JAK complex and inhibits its kinase activity (step 10). In this pathway, SHP-2, PPX, and PPN work as phosphotases and try to inhibit activity of phosphorylated proteins [8].

III. Simulation and Results

A. Mathematical Model

To evaluate the dynamics of JAK/STAT signaling pathway, the model and kinetic parameters were adopted from [8]. Representing all reactions by mass-action kinetics, an ODE model is formulated for pathway. As an example, consider a simple binding reaction, $\mathrm{R}+\mathrm{J} \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} \mathrm{RJ}$, where receptor R binds to ligand J and complex RJ is formed. k_1 and k_{-1} represent forward and backward rate constant respectively. Rate equation for preceding reaction can be written as following:

$$v = k_1 * [R] * [J] - k_{-1} * [RJ] \tag{1}$$

and the dynamics of the system can be described by differential equation as following

$$\frac{d[R]}{dt} = -v, \quad \frac{d[J]}{dt} = -v, \quad \frac{d[RJ]}{dt} = v \tag{2}$$

Thus, the ODEs describing the dynamics of all chemical species within the pathway can be constructed. In this work, we are interested in time-course of $STAT1n^* - STAT1n^*$, which acts as a transcription factor inside the nucleus as shown in Fig. 2. Double hump pattern in Fig. 2, verifies the inhibition of $STAT1n^* - STAT1n^*$ by induced $SOCS1$.

B. Sensitivity Analysis

The Sobol sensitivity analysis was performed based on the numerical approach in [6]. At first step, two sets including N sample values for each input variable is generated following an uniform distribution for creating sample set. The range of parameter distribution is usually determined from available literature [5]. These two sets are called sample and resample sets. Given the number of parameters, p, Sobol indices can be computed in $p+2$ model runs. Each run consists of N iterations. Let f represent the vector of model output, then the main (S_j) and total Sobol index (T_j) is calculated using following formulas:

$$S_j = \frac{u_j - m_0'^2}{v_0}, \quad T_j = 1 - \frac{u_{-j} - m_0^2}{v_0} \tag{3}$$

where

$$u_j = \frac{1}{N}\sum f_0' f_j, \quad u_{-j} = \frac{1}{N}\sum f_0 f_j, \quad v_0 = \frac{1}{N}\sum f_0^2 - m_0^2$$
$$m_0 = \frac{1}{N}\sum f_0, \quad m_0' = \frac{1}{N}\sum f_0' \tag{4}$$

f_j is the output from the run that used resample set for parameter j and sample set for all other parameters, and

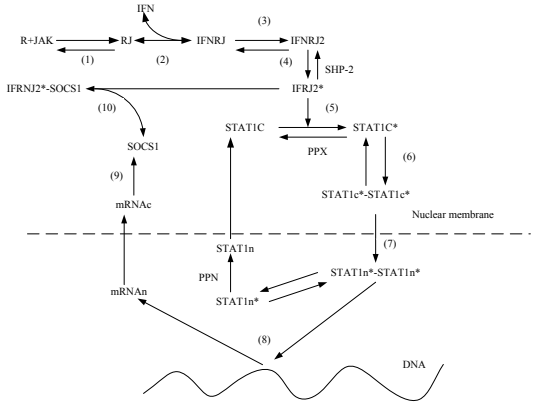

Fig. 1. Kinetic model of JAK/STAT signaling pathway

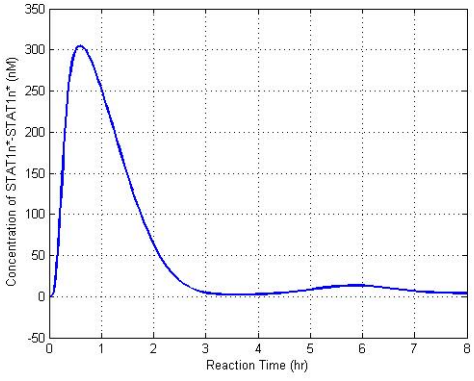

Fig. 2. Simulated time course of $STAT1n^* - STAT1n^*$ activated by continuous exposure to IFN-γ

TABLE I

RESULT OF SOBOL SENSITIVITY ANALYSIS WITH RESPECT TO VARIATION IN INITIAL CONCENTRATION OF METABOLITES

Component	Initial Value(nM)	Range of Variation	Sobol value(Total effect)	Sobol value(Main effect)
IFN	4	0.4-40	0.5504	0.3562
PPN	60	6-600	0.4398	0.2647
STAT1C	1000	100-10000	0.2581	0.1107
PPX	50	5-500	0.1081	0.0534
JAK	12	1.2-120	0.0691	0.0324
R	12	1.2-120	0.0549	0.0306
SHP-2	100	10-1000	0.0498	0.0351

f_0' is the output from the run using resample set for all parameters, also f_0 is the output produced considering the sample set for all parameters. The number of samples in our simulation was set to 2000. In order to obtain precise estimates for Sobol index, the entire analysis is repeated. This is a well-known technique called bootstrapping [6]. The reference values for the initial concentrations of component and variation ranges used for simulation are shown in Table I. The table entries are sorted according to their total Sobol values. Main effect and total effect for these components are depicted in Fig. 3. In this work we are interested in studying the effect of variation in initial concentration of the phosphotases(SHP-2, PPN, and PPX), on the activation of transcription factor($STAT1n^* - STAT1n^*$). As the results shows, PPN has the most dominant effect, based on its total Sobol index.

C. Optimal Control

The behavior of the ODE model describing the pathway, can be affected by variables which can be controlled from the outside the model. The question arises as to how to control the external components in order to produce the "best" outcome or guide the pathway through a desired set of states, as measured by some predetermined goals. The mathematical theory that answers this question is called optimal control theory. This theory is a powerful mathematical tool that can be used to make decision involving complex biological systems. For example, minimize the growth of a certain harmful bacterial population while keeping the level of the toxic drugs low [9]. The ODE model of any biochemical pathway describing its dynamics can in general be represented as a state space equations as following:

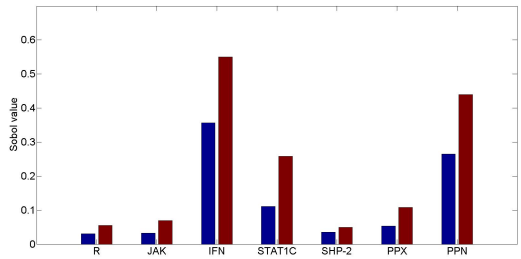

Fig. 3. Results of Sobol sensitivity analysis ,blue and red bars represent main effect and total effect respectively

$$\dot{X} = g(t, X(t), U(t)) \tag{5}$$

where $U(t)$ and $X(t)$ represent control input and state respectively. Basic optimal control problem consists of finding a piecewise continuous control $U(t)$ and the associated state variable $x(t)$ to maximize or minimize the following objective function over control input $u(t)$ over the time interval $[t_0, t_f]$.

$$J(x(t), u(t), t) = \int_{t_0}^{t_f} f(x(t), u(t), t)dt \tag{6}$$

Subject to:

$$\dot{X} = g(t, X(t), U(t)),\ t_0 \leq t \leq t_f \tag{7}$$

$$X(t_0) = X_0,\ \textit{and}\ X(t_f)\ \textit{is free} \tag{8}$$

The principle technique for solving preceding problem is Pontryagin maximum principle which provides a set of "necessary conditions" that an optimal control and corresponding state must satisfy. More detail about these procedure can be found in [10]. Overexpression and constitutive activation of STAT proteins occurs in many ailments including prostate cancer [11] and inflammatory bowel diseases (IBD) [12]. In this work, we are interested in controlling the activation process of STAT protein inside nucleus, so the goal of optimization is to identify a regulatory process in the form of an external control input or external stimulus, which keeps the concentration of $STAT1n^*$ dimer within a certain range. To this end, we try to find a suitable control input to minimize the following objective function

$$J = \frac{1}{T}\int_0^T ([STAT1n^* - STAT1n^*]^2 + \frac{1}{2}u^2(t))\mathrm{d}t \tag{9}$$

where T is the time interval over which we are interested in studying the behavior of system. The complex dynamics of biological pathway consists of several nodes and inputs any of which can be used for injection of a control input. The sensitivity analysis described in the previous section can be extremely useful in determining which input nodes can be used for control. Hence, we analyze the cost function by injecting external control input to dynamics of 3 phosphotases(PPN, SHP-2, and PPX) separately. The optimal control

for each of these inputs considered separately can be synthesized via the technique of steepest descent algorithm[10]. The results depicted in Fig. 4 shows that by applying external input with higher Sobol index, better value for cost function is attained.

IV. Discussion

The results above, strongly support the claim that the nodes with highest Sobol sensitivity indices can be selected in order to synthesize inputs that yields a more cost-effective control. As shown in Fig. 4, the components that the output of the pathway bears a greater sensitivity towards, can be used to produce a more desired(cost-effective) behavior of the output. The mathematical model combined with sensitivity analysis can be used to identify key factors including dominant components as well as critical reactions in signaling pathway [13]. The proposed approach in this work is based on developing a regulatory process to obtain a desirable behavior from system, where we defined the cost as a function of control input and concentration of target component ($STAT1n^* - STAT1n^*$). In this study we show that, results from sensitivity analysis can be extremely useful to synthesize optimal control.

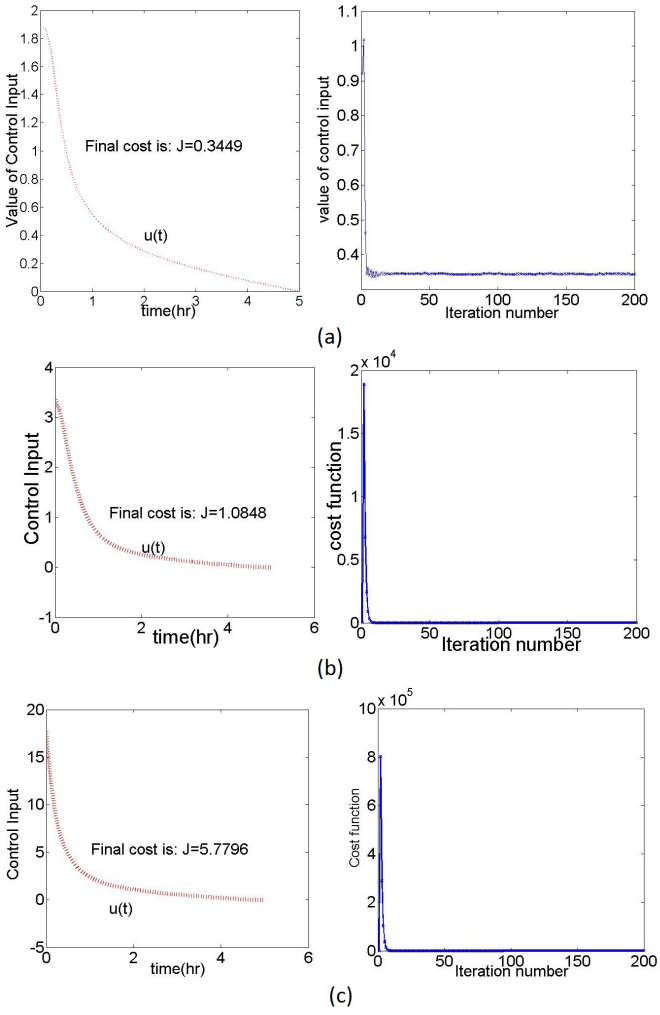

Fig. 4. Top-down, the cost of heuristic optimal control for PPN, PPX, SHP-2 respectively which bears strong correlation to their respective sensitivity indices. The panels on the right indicate the best-cost vs iteration for the optimization procedure.

V. Conclusions

JAK mutations in humans can cause numerous diseases, including severe combined immune deficiency, leukemias, cancer, and other myeloproliferative disorders. Because of the causative role in these diseases and their central significance in immune response, JAK/STAT pathways have become attractive targets for development of therapeutics for a variety of immune system disorders. Developments over the past decades in biological pathway modeling have led to an increasing complexity and number of the available tools. In this work we have shown that, the Sensitivity Analysis(SA) can be effectively leveraged to synthesize optimal control via Pontryagin's maximum principle that helps design cures in time and cost-effective fashion. This could have broader impact in treatment of medical ailments and suppress symptoms. This information can also be useful in gaining a deeper insight and hypothesize important mechanism in a system, as well as within a "control" context, suggesting how we may intervene in the system to produce certain behavior.

References

[1] H. De Jong, "Modeling and simulation of genetic regulatory systems: a literature review," *Journal of computational biology*, vol. 9, no. 1, pp. 67–103, 2002.

[2] G. Archer, A. Saltelli, and I. Sobol, "Sensitivity measures, anova-like techniques and the use of bootstrap," *Journal of Statistical Computation and Simulation*, vol. 58, no. 2, pp. 99–120, 1997.

[3] I. M. Sobol, "Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates," *Mathematics and computers in simulation*, vol. 55, no. 1-3, pp. 271–280, 2001.

[4] Y. Zheng and A. Rundell, "Comparative study of parameter sensitivity analyses of the tcr-activated erk-mapk signalling pathway," *IEE Proceedings-Systems Biology*, vol. 153, no. 4, pp. 201–211, 2006.

[5] J. Nossent, P. Elsen, and W. Bauwens, "Sobolsensitivity analysis of a complex environmental model," *Environmental Modelling & Software*, vol. 26, no. 12, pp. 1515–1525, 2011.

[6] G. Glen and K. Isaacs, "Estimating sobol sensitivity indices using correlations," *Environmental Modelling & Software*, vol. 37, pp. 157–166, 2012.

[7] F. Wessely, M. Bartl, R. Guthke, P. Li, S. Schuster, and C. Kaleta, "Optimal regulatory strategies for metabolic pathways in escherichia coli depending on protein costs," *Molecular systems biology*, vol. 7, no. 1, 2011.

[8] S. Yamada, S. Shiono, A. Joo, and A. Yoshimura, "Control mechanism of jak/stat signal transduction pathway," *FEBS letters*, vol. 534, no. 1, pp. 190–196, 2003.

[9] S. Lenhart and J. T. Workman, *Optimal control applied to biological models*. CRC Press, 2007.

[10] D. E. Kirk, *Optimal control theory: an introduction*. Courier Dover Publications, 2012.

[11] P. Kroon, P. A. Berry, M. J. Stower, G. Rodrigues, V. M. Mann, M. Simms, D. Bhasin, S. Chettiar, C. Li, P.-K. Li *et al.*, "Jak-stat blockade inhibits tumor initiation and clonogenic recovery of prostate cancer stem-like cells," *Cancer research*, vol. 73, no. 16, pp. 5288–5298, 2013.

[12] R. Atreya and M. F. Neurath, "The role of the jak/stat pathway in inflammatory bowel diseases," *JAK-STAT Pathway in Disease*, p. 90, 2009.

[13] Z. Zi, K.-H. Cho, M.-H. Sung, X. Xia, J. Zheng, and Z. Sun, "In silico identification of the key components and steps in ifn-γ induced jak-stat signaling pathway," *FEBS letters*, vol. 579, no. 5, pp. 1101–1108, 2005.