Optimization-based Inference for Temporally Evolving Boolean Networks with Applications in Biology

Young-Hwan Chang
Joe Gray
Claire Tomlin

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2010-133
http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-133.html

October 26, 2010

Acknowledgement

The first author would like to thank support from STX foundation from Korea. This research was supported by the NIH NCI through the ICBP and PS-OC projects. This paper was submitted for consideration as a regular paper at the 2011 American Control Conference, in September 2010.

Optimization-based Inference for Temporally Evolving Boolean Networks with Applications in Biology

Young Hwan Chang, Joe Gray and Claire Tomlin

***Abstract*—Modeling of biological genetic networks forms the basis of systems biology. In this paper, we present an optimization-based inference scheme to identify temporally evolving Boolean network representations of genetic networks from data. In the formulation of the optimization problem, we use an adjacency map as *a priori* information, and define a cost function which both drives the connectivity of the graph to match biological data as well as generates a sparse and robust network at corresponding time intervals. Throughout simulation studies on simple examples, it is shown that this optimization scheme can help to understand the structure and dynamics of biological genetic networks.**

I. Introduction

MODELING of biological genetic networks has received much recent research attention. Much work has been done to develop Bayesian network models of genetic networks by coding *a priori* knowledge on the regulatory relationships into probabilistic models [1][2][3]. On the other hand, there are many studies of identification of regulatory networks using deterministic models such as ordinary differential equations (ODEs) or linear models based on least squares identification [4][5][6]. Also, there are many applications of Boolean networks to modeling and analyzing biological systems, as well as an increase of research activities to address questions arising from biological applications [7][8]. Often, though, models of biological systems are too complex to understand because of the large number of components involved and the nonlinearity of the reaction or interaction. As a result, the behavior of these systems in general cannot be completely understood from a systems point of view. Moreover, once a model structure is chosen, such as a mass kinetic reaction model or a nonlinear ODE model, prejudices of the model are automatically imposed which then restrict the representation and understanding of biological data.

Since a graph is a natural way to represent a biological network, if a system can be abstracted into a graph, it can help to understand the biological network. A graph is a set of vertices which represents states, and a set of edges which depicts the relationship or connection between two or more states. A given connectivity or adjacency map is a signed, directed graph $GR = (V, E, S)$ where V is a set of vertices, E is a set of directed edges, and $S : E \to \{-1, 0, +1\}$. For example, $e_{ij} = 1$ represents the case in which input node j activates output node i. If input node j inhibits output node i, then $e_{ij} = -1$. If input node j does not affect output node i, then $e_{ij} = 0$. Also, graphs are well-suited for situations in which there is little prior or explicit knowledge of the dynamics. Moreover, if we can build a graph model to represent biological data, we could escape imposed prejudices from the model structure. There are several graph mining approaches to biological networks [9][10]. These approaches represent biological networks as graphs, where nodes represent genes and edges represent relationships between each gene, and discover frequent patterns or *motifs* [9] in these graphs. These approaches focus on structural features of networks and they can effectively uncover the functional interaction structure of a biological network. Also, these approaches consider time-invariant networks and local or modular behavior of large networks. Recent studies [11][12] have proposed a concept of a "temporal sequence of network motifs" where the motifs change according to the dynamic nature of the biological system and can describe pivotal developmental events which cannot be captured by the static network approach: the former [11] develops algorithms for graph-rewriting rules based on machine learning techniques, which brings complexity issues which analyzing very large graphs [11]. On the other hand, [12] applies a temporal sequence of network motifs analysis by reconstructing the active sub-networks (3-node sub-graphs).

The main idea of our scheme lies in representing the captured relationship as a network path with *a priori* information (a given connectivity map) and using convex optimization techniques to find the time-varying sparsest graph consistent with experimental observations. Despite uncertainties about details for a given biological system, we often have reasonable qualitative knowledge about interactions of each gene, so we can use this information as *a priori* information. In this setting, the model behavior is solely based on this qualitative information which guarantees biologically reasonable behavior: robustness and sparsity in general. The ability of many biological networks to exhibit their function reliably despite noise or perturbation is often referred as functional robustness. We also note that biological regulatory networks are likely to be sparse especially at a fixed interval of time (for example, most transcription factors (TFs) do not regulate most genes). Also, there are expectations behind modeling efforts:

Y.H. Chang is with the Department of Mechanical Engineering, University of California, Berkeley, CA 94720 USA (e-mail: yhchang@berkeley.edu)

Prof. Joe Gray is the Director, Life Sciences Division, Lawrence Berkeley National Laboratory, and Adjunct Professor of Laboratory Medicine, UCSF (e-mail: jwgray@lbl.gov)

Prof. Claire Tomlin is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720 USA (corresponding author, e-mail: tomlin@eecs.berkeley.edu).

1) Networks represent the structure of complex connections so viewing evolving networks as dynamical systems allows us to predict many of their properties analytically.
2) If we can match signal propagation that drives the placement of links and nodes, then the topology or the structural elements will follow. It can help to move beyond architecture and uncover the laws that govern the underlying dynamic process.

In contrast to previous methodologies for dynamic graph analysis [11][12], in this paper, we develop a convex optimization-based inference method, where we embed the dynamics of a linear time varying representation, and enforce sparsity at corresponding time intervals. The rest of this paper is organized as follows: Section II presents an overview related to modeling of biological networks. An optimization problem formulation is discussed in Section III. Simple examples are given in Section IV. Also, Section V presents an example of biological network of HER2 over-expressed breast cancer. Finally, conclusions are given in Section VI.

II. Overview

We define a state vector $x(t) = [x_1(t), ..., x_{n_x}(t)]^T$, the components of which represent concentrations of proteins or states in a biological network. The evolution of state $x(t)$ can be modeled using an ordinary differential equation (ODE):

$$\dot{x}(t) = f(x(t), p) \tag{1}$$

where p is a parameter set. The nonlinear dynamic system (1) can be approximated by a linear system based on forming the Jacobian around steady states as shown below:

$$\delta\dot{x}(t) = \frac{\partial f}{\partial x}\delta x(t) + \frac{\partial f}{\partial p}\delta p = A\delta x(t) + B\delta p \tag{2}$$

A system in the form of (2) can be considered as a weighted directed graph. Then, A represents connectivity and B represents the sensitivity of parameter variation. If A_{ij} is zero, node j has no direct effect on node i. Also, if $A_{ij} > 0$, node j activates node i. Similarly, if $A_{ij} < 0$, node j inhibits node i. In [10], a convex optimization is constructed as follows:

$$\begin{aligned} &\min_{A,\bar{B}} \left\| (\dot{\tilde{X}} - \bar{B} - A\tilde{X})W \right\|_F \\ &\text{subject to } card(A) \le k, A_{i,j} > 0, A_{r,s} < 0 \end{aligned} \tag{3}$$

where $\tilde{X}(= [X_1\ X_2\ ...\ X_L])$ represents the time course data set with different stimulations and/or inhibitions and each X_i represents the matrix form of n_x different components at M different time points $X_i = \begin{bmatrix} x^i_{1,1} & x^i_{1,2} & ... & x^i_{1,M} \\ x^i_{2,1} & x^i_{2,2} & ... & x^i_{1,M} \\ ... & ... & ... & ... \\ x^i_{n_x,1} & x^i_{n_x,2} & ... & x^i_{n_x,M} \end{bmatrix}$. Also, $\bar{B}\,(= [B_1\ B_2\ ...\ B_L])$ represents the set of sensitivities of parameter variation with $B_i = \overbrace{[b_i\ ...\ b_i]}^{M}$ and Wrepresents a weighting matrix for specific experiments. Also, k is a given positive constant which represents maximum connectivity, all $A_{i,j} > 0$ represent activation edges (node j activates node i) and all $A_{r,s} < 0$ represent inhibition edges (node j inhibits node i). Therefore, this approach gives us the optimal static graph map consistent with various experimental data sets.

In this paper, we extend this idea to a dynamic graph model. First, we define $\mathcal{X} = [X_N^T, X_{N-1}^T, ..., X_1^T]^T \in \mathbf{R}^{n_x \cdot N \times 1}$ where $X_k \in \mathbf{R}^{n_x \times 1}$ is a snapshot of data or known vector (normalized or Booleanized biological data) at time k for $1 \le k \le N$, the components of which represent concentrations or activities in a biological network where n_x is a number of states of X_k and N is a number of discrete time steps. We define $\mathcal{G} = \mathcal{F}(G_1, G_2, ..., G_N)$ as an augmented matrix of dynamic graph G_k where each $G_k \in \mathcal{R}^{n_x \times n_x}$ is a connectivity map at time k for $1 \le k \le N$ which is based on *a priori* information GR. Also, the augmented matrix, $\mathcal{G}$, satisfies an evolution of the state X_k with time and follows the linear relation $X_k = G_k X_{k-1}$. Moreover, we define *a common* sub-graph S_k, *an addition* sub-graph A_k and *a removal* sub-graph R_k at time k as follows [11]:

$$\begin{aligned} S_k &= G_k \cap G_{k-1} \\ A_k &= G_k \cap S_k^c \\ R_k &= G_k^c \cap G_{k-1} \end{aligned} \tag{4}$$

where superscript '*c*' indicates the complement of a set in its ambient space. Therefore, S_k represents common sub-graph between G_k and G_{k-1} and (A_k, R_k) shows the evolution of the graph at time k indicating how the biological system changes over time.

III. Optimization Problem Formulation

In contrast to previous methodologies for dynamic graph analysis [11][12], in this section, we formulate a convex optimization-based inference method, where we embed the dynamics of a linear time varying representation, and enforce sparsity and smooth evolution at corresponding time intervals.

A. Dynamic Graph (Linear Time Varying System)

The state X_k evolves along with time and constitutes the following linear time varying system:

$$X_k = G_k X_{k-1} \tag{5}$$

where $G_k = g(GR, X_k|X_{k-1}) = h(C_k, W_k)$, where C_k represents connectivity and W_k is a weighting factor or strength of connection. Therefore, G_k is a connectivity map at time k which is based on *a priori* knowledge map (GR) and consistent with experimental data. Note that G_k describes how the evolution of components of X_k depends on interactions with X_{k-1} based on GR. For example, we have given candidate edges in GR, only a few of these may be selected, based on the relationship between X_k and X_{k-1}. If all the interactions between each component are properly identified, we can reconstruct the function G_k in terms of the connectivity and weighting factor. For instance, $G_k(i, j) = 0.5$ represents that node j activates node i with strength 0.5. The strength is related to the reaction rate and the concentration of other species, demonstrated by the Jacobian of a mass kinetic reaction model.

The goal of system identification of biological systems is to infer each G_k for $1 \leq k \leq N$ consistent with both biological data set $\mathcal{X}$ and *a priori* information GR. In general, a gene regulation network (GRN) has the following characteristics [13]:

1) Directionality: regulatory control is directed from regulators to regulated genes.

2) Sparsity: each single gene is controlled by a limited number of other genes, which is small compared to the total gene content (and also to the total number of TFs) of an organism.

3) Combinatorial control: the expression of a gene may depend on the joint activity of various regulatory proteins.

Since GRNs have a sparse structure with combinatorial control, we should reconstruct the sparsest graph consistent with experimental observations. We can construct an optimization problem as follows:

$$\begin{aligned} &\min_{G_k} \|X_k - G_k X_{k-1}\| + \gamma \|A_k\| \\ &\text{subject to} \quad G_k = g(GR, X_k|X_{k-1}) \end{aligned} \tag{6}$$

where the second term in the cost function penalizes the cost of adding edges in order to avoid heavy combinatoric computation. This term therefore enforces the network to be sparse, where γ is a positive constant, representing the trade-off between reconstruction error and sparsity. Here, we define a function g as shown below:

$$G_k = g(GR, X_k|X_{k-1}) = GR \oplus MAP = \mathsf{Proj}_{MAP} GR \tag{7}$$

where $\oplus$ is defined as a projection operator onto $MAP \in \mathbf{R}^{n_x \times n_x}$ whose i-th column is a column vector, the components of which are all one if $X_{k-1}(i)$ is active, which means the state of the i-th element is one or over the threshold. On the other hand, if $X_{k-1}(i)$ is non-active, then the i-th column of MAP is a zero column vector. Therefore, basically, this projection gives us all possible candidate edges based on both X_{k-1} and GR. For example, if x_i at the $(k-1)$*th* step is active, then i-th column of GR contains the candidate edges. On the other hand, if x_i at $(k-1)$*th* step is not active, we cannot use the i-th column of GR as candidate edges.

If we implement the optimization problem step by step independently, the penalty term for sparsity does not play the role of generating sparse network, but, instead distributing power to the dense network which uses all possible edges because the dense network with distributed power gives us the lower cost according to this formulation. However, if we formulate the problem considering the overall time step, then the optimal solution gives us the sparsest and smooth evolving network which selects corresponding effective edges. Also, we can reformulate sub-graph A_k as follows:

$$A_k = G_k \cap S_k^c = G_k \cap (G_k \cap G_{k-1})^c = G_k - G_{k-1} \tag{8}$$

Then, we can construct a convex optimization problem for the proposed identification problem as shown below:

$$\begin{aligned} \min_{G_1,\ldots,G_N} &\sum_{k=1}^{N} \|X_k - G_k X_{k-1}\|_2 + \gamma\{\sum_{k=2}^{N} \|G_k - G_{k-1}\|_F \\ &+ \|G_1\|_F + \|G_N\|_F\} \\ \text{subject to} \quad & G_k = g(GR, X_k|X_{k-1}) \end{aligned} \tag{9}$$

Note that the first term of equation (9), the summation of $\|X_k - G_k X_{k-1}\|$ forces to minimize the reconstruction error for a given dynamical network at time k for $1 \leq k \leq N$. Also, we consider the second term, the summation of $\|G_k - G_{k-1}\|_F$ which plays the role of realizing a smooth evolution and minimizes the change of network evolution. Finally, with the penalty term $\|G_1\|_F + \|G_N\|_F$ which plays as a boundary constraint, we can find the optimal sparsest dynamic graph. We can also arrange and reformulate equation (9) as follows:

$$\begin{aligned} \min_{\mathcal{G}} &\|\mathcal{X} - \mathcal{G}\mathcal{X}\|_2 + \gamma \left\|(\mathcal{G}^T - \mathcal{G}) \times W\right\|_F \\ &\text{subject to} \quad \text{given } \mathcal{X}, W \\ &\quad G_k^{act} \geq 0 \\ &\quad G_k^{inhib} \leq 0 \\ &\quad G_k^{others} = 0 \end{aligned} \tag{10}$$

where

$$\mathcal{X} = \begin{bmatrix} X_N \\ X_{N-1} \\ \cdots \\ X_1 \\ X_0 \end{bmatrix}, \mathcal{G} = \begin{bmatrix} O_{n_x} & G_N & O_{n_x} & \cdots & O_{n_x} \\ O_{n_x} & O_{n_x} & G_{N-1} & \cdots & O_{n_x} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O_{n_x} & O_{n_x} & \cdots & O_{n_x} & G_1 \\ O_{n_x} & O_{n_x} & \cdots & O_{n_x} & I_{n_x} \end{bmatrix}$$

and $W = \begin{bmatrix} I_{n_x} & O_{n_x} \\ O_{n_x} & I_{n_x} \\ I_{n_x} & O_{n_x} \\ O_{n_x} & I_{n_x} \\ \cdots & \cdots \\ \cdots & \cdots \end{bmatrix}$ where $\mathcal{X} \in \mathbf{R}^{n_x \cdot (N+1) \times 1}$,

$\mathcal{G} \in \mathbf{R}^{n_x \cdot (N+1) \times n_x \cdot (N+1)}$ and $W \in \mathbf{R}^{n_x \cdot (N+1) \times 2 \cdot n_x}$ for $1 \leq k \leq N$. Note that the first term of the cost function in equation (10) is a reconstruction error cost and the second term plays the role of realizing a smooth and sparse evolution of the network by selecting effective edges.

B. Static Graph (Linear Time Invariant System)

If we assume that the graph model does not evolve with time (static graph) such as with a linear time invariant system [10], we can modify the structure $\mathcal{G}$ and constraint as shown below for a fixed pattern graph:

$$\mathcal{G} = \begin{bmatrix} O_{n_x} & G & O_{n_x} & ... & O_{n_x} \\ O_{n_x} & O_{n_x} & G & ... & O_{n_x} \\ ... & ... & ... & ... & ... \\ O_{n_x} & O_{n_x} & O_{n_x} & O_{n_x} & G \\ O_{n_x} & O_{n_x} & O_{n_x} & O_{n_x} & I_{n_x} \end{bmatrix}$$
$$G^{act} \geq 0 \ , \ G^{inhib} \leq 0, \ G^{others} = 0 \tag{11}$$

where $G = \tilde{g}(GR)$ does not depend on time (compared with $G_k = g(GR, X_k | X_{k-1})$ for a linear time varying system). Note that for a fixed pattern, the optimal solution represents the average connectivity map.

C. Dynamic and Static Graph

We can compare the dynamic graph and static graph method: the main difference of cost function from dynamic and static graph is the penalty for sparsity as follows:

$$\left\| (\mathcal{G}^T - \mathcal{G}) W) \right\|_{F,dynamic} = \begin{bmatrix} O_{n_x} & -G_N \\ \Delta G_{N-1} & O_{n_x} \\ O_{n_x} & \Delta G_{N-2} \\ ... & ... \\ O_{n_x} & \Delta G_1 \\ G_1 & O_{n_x} \end{bmatrix} \tag{12}$$

$$\left\| (\mathcal{G}^T - \mathcal{G}) W) \right\|_{F,static} = \begin{bmatrix} O_{n_x} & -G \\ O_{n_x} & O_{n_x} \\ O_{n_x} & O_{n_x} \\ ... & ... \\ O_{n_x} & O_{n_x} \\ G & O_{n_x} \end{bmatrix} \tag{13}$$

where $\Delta G_k = G_{k+1} - G_k$. Also, if we modify the constraint for a dynamic graph similar to static graph approach, a dynamic graph approach gives us a lower cost than the static graph approach because the structural constraint restricts the degree of freedom of choosing edges:

$$\min_{\mathcal{G}} \left\| \mathcal{X} - \mathcal{G}\mathcal{X} \right\|_2 + \gamma \left\| (\mathcal{G}^T - \mathcal{G}) \times W \right\|_F \leq \min_{\bar{\mathcal{G}}} \left\| \mathcal{X} - \bar{\mathcal{G}}\mathcal{X} \right\|_2 + \gamma \left\| (\bar{\mathcal{G}}^T - \bar{\mathcal{G}}) \times W \right\|_F \tag{14}$$

where $\mathcal{G}$ represents the optimal solution of the dynamic graph approach and $\bar{\mathcal{G}}$ represents the optimal solution of the static graph approach.

D. Inhibition Edges

Based on our formulation of the optimization problem, we can find the optimal solution which satisfies a trade-off between representation of data (dynamics) and sparsity. However, the optimal solution does not include any inhibition edges because it is not necessary as shown in Figure 1. For example, if "X" is active (1) and "Y" is not active (0), then there are two possible cases: X inhibits Y (connected, Fig 1 (left)) or no connection between X and Y (Fig 1 (right)):

Figure 1: Possible cases for inhibition edge where 1 represents activated gene and 0 represents deactivated gene.

However, we can handle inhibition edges using Boolean logic as an algebraic constraint as shown below:

$$\mathsf{Y} = \mathsf{not}\ \bar{\mathsf{Y}}\ (=\sim \bar{\mathsf{Y}}) \tag{15}$$

Also, we extend this algebraic constraint for normalized state as shown below:

$$\mathsf{Y} + \bar{\mathsf{Y}} = 1 \tag{16}$$

Consider the simple case shown in Figure 2. Basically, state X inhibits state Y. Using an algebraic constraint (15), we can change the inhibition edge to an activation edge with the new state $\bar{\mathsf{Y}}$ as shown below:

$$\mathsf{X} \dashv \mathsf{Y} \implies \mathsf{X} \dashv (\sim \bar{\mathsf{Y}}) \implies \mathsf{X} \rightarrow \bar{\mathsf{Y}} - \mathsf{Y} \tag{17}$$

Hence, we extend states if there are inhibition input edges and introduce a diagonal weighting matrix $\mathcal{M}$, which makes all species have the same penalty as shown below:

$$\min_{\mathcal{G}} \left\| \mathcal{M} \times (\bar{\mathcal{X}} - \mathcal{G}\bar{\mathcal{X}}) \right\|_2 + \gamma \left\| (\mathcal{G}^T - \mathcal{G}) \times W \right\|_F \tag{18}$$

where $\bar{\mathcal{X}}$ represents extended states and $\mathcal{M}_{ii} = \{1, \frac{1}{\sqrt{2}}\}$. If there exist $x, \bar{x}$ for a specific state, $\mathcal{M}_{ii} = \frac{1}{\sqrt{2}}$ and otherwise, $\mathcal{M}_{ii} = 1$.

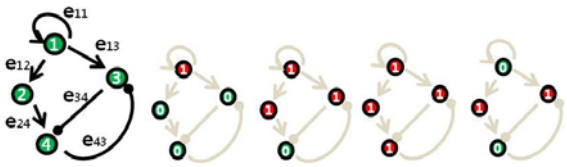
Figure 2: (A) Inhibition edge (X inhibits Y) (B) modified edge (X activates $\bar{\mathsf{Y}}$)

IV. Examples

In this section, we consider toy examples to illustrate the proposed inference scheme.

A. *Simple Gene Network*

We first consider a toy example composed of four genes. *A priori* information and the snapshot of gene expression are shown in Figure 3. Here we do not consider state extension for inhibition edges which means the optimal solution does not include any inhibition edges.



(a) (b)

Figure 3: (a) *a priori* connectivity map where the arrows indicate activation and blunted lines denote inhibition. (b) snapshots of gene expression from time k=1 to k=4 (red or 1: activated states, green or 0: deactivated states).

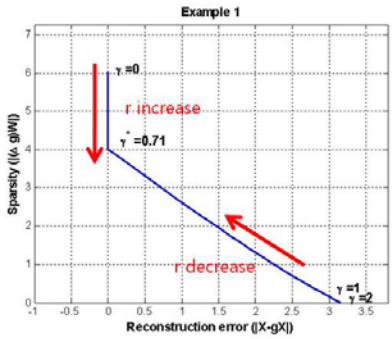


Figure 4: Trade-off curve between the model fitting and the sparsity with varying parameter γ.

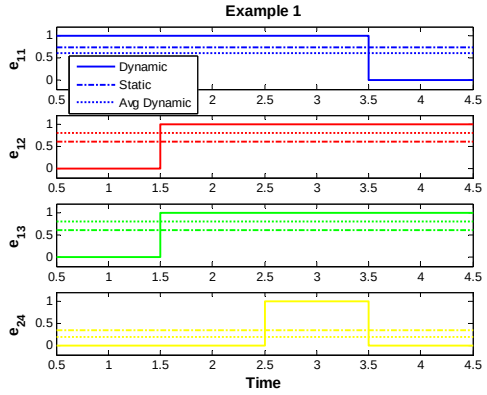
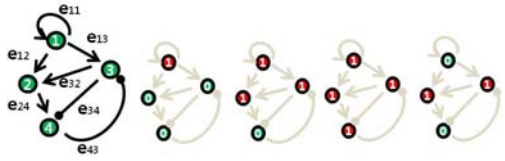


Figure 5: The optimal solution for example A: the magnitude of each edge represents strength of connection.

By varying parameter γ, we can sweep out the optimal trade-off curve between the reconstruction error and the sparsity of a solution as shown in Figure 4. We can choose the optimal parameter γ^* by the graphical representation: the extreme point γ^* on the trade-off between the sparsity and the reconstruction error. Once we fix the parameter γ^*, we solve the constrained convex optimization problem (10) using CVX[14]. Figure 5 shows the dynamics of the connectivity graph. We can capture the temporal graph not only in terms of connection but also strength of the edge. Also we can compare two approaches: dynamic and static graph approach with average of dynamic graph.

B. *Simple Gene Network with different structure*

Here, we add 1 edge which connects vertex 3 to vertex 2 as shown in Figure 6 and solve the optimization problem again. We can see the difference of the strength of edge 12 (e_{12}) compared with example A. Basically, for example A, the optimal graph shows the robust pathway distributing power evenly (e_{12} and e_{13} in Figure 5) because both pathways are effective equally. However, for example B, an additional pathway (e_{32}) changes the topology of graph which makes the optimal graph to choose the more effective or economical path ($e_{13} - e_{32}$), which represents the selectivity. In other words, the optimal solution shows that the strength of e_{12} decreases because there exists a more effective pathway ($e_{13} - e_{32}$).



(a) (b)

Figure 6: An additional edge which connects from node 3 to node 2.

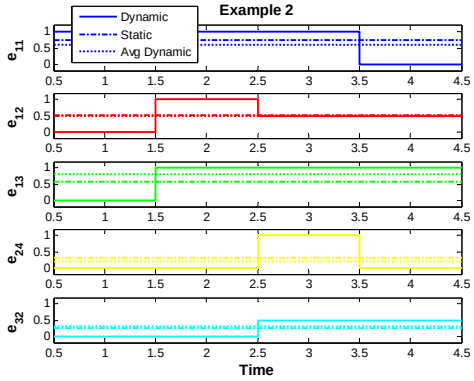


Figure 7: An optimal solution for example B.

C. *p53 Signal Pathway*

Aswani et al [15] proposed a graph-theoretic topological control applied to the p53 signal pathway. We apply our approach to understand how the controller affects the biological pathway and capture the evolution of signal pathway. We define $X = [x_1, x_2, x_3, x_4]^T = [x_{MDM}, x_{p53}$ $x_{cyclinG}, x_c]^T$ where x_c is a virtual state which represents the proposed control scheme (actually removing the edge in [15]):

$$\begin{aligned} x_c &= x_{MDM} \text{ if controller Off} \\ x_c &= 0 \qquad \text{if controller On} \end{aligned} \tag{19}$$

Hence, by introducing this virtual state, we have an abstract model of abnormal p53 signal pathway with controller in Figure 8. Also, we can define GR as follows based on Figure 8 with extending states including the state extension due to incorporating the inhibition edges:

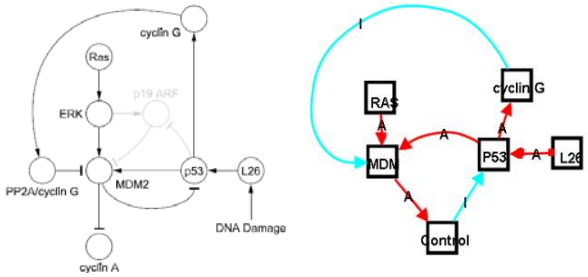


Figure 8: (left) an abnormal p53 pathway in Figure 3 (c)[15] (right) the abstract model which includes the effect of controller where A (red) represents activation edge and I (blue) represents inhibition edge.

$$GR = \overbrace{\begin{matrix} x_1 & \bar{x}_1 & x_2 & \bar{x}_2 & x_3 & x_4 \end{matrix}}$$
$$\begin{bmatrix} 0 & 0 & e_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & e_{3\bar{1}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & e_{4\bar{2}} \\ 0 & 0 & e_{23} & 0 & 0 & 0 \\ e_{14} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (20)$$

Here, we normalize the data then apply our algorithm. We can capture the dynamic evolution of graph in Figure 9. The controller causes the p53 concentrations to increase to higher levels by regulation edge from MDM2 to p53 and causes increased strength of inhibition edge ([p53]-[cyclin G]-[MDM2]). In other words, p53 regulates MDM2 similar to the normal p53 pathway [15]. If the controller is not applied again, the strength of edge ([p53]-[cycle G]-[MDM2]) decreases and the strength of activation edge [p53]-[Controller]-[MDM2] increases. This causes MDM concentrations to increase to higher levels which cause regulation p53 by inhibition edge ([MDM2]-[Controller]-[p53]) similar to the abnormal p53 pathway [15].

We can also apply our algorithm for the normal p53 pathway in order to compare with the abnormal p53 pathway with controller. In a normal p53 pathway, we can consider all possible combinations of both Ras and L26 as two input signals. Here, the basic assumption is that the inhibition reaction is stronger than the activation reaction. Then, we found that the p19 ARF mainly regulates MDM2 and it can not affect MDM2 from p53 through p19 ARF as shown in Figure 10. Hence, we can use the same abstract model in Figure 8 for a normal p53 signal pathway. The optimal solution gives us that the normal p53 cell uses mainly inhibition edges from p53 to MDM2 through cyclin G which means p53 regulates MDM2. Therefore, the controller drives the abnormal p53 cell to the normal p53 cell by removing the inhibition edge from MDM2 to p53 as Aswani et al [15] proposed.

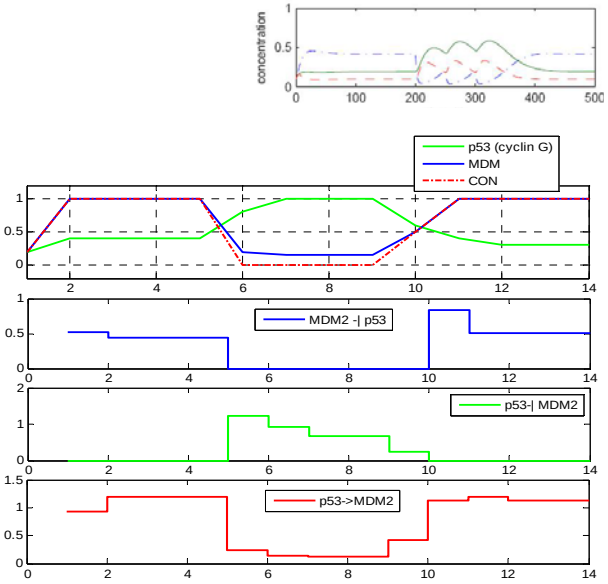


Figure 9: (upper) time course plots for the abnormal p53 pathway with controller in Figure 4 (c) [15] (lower) dynamic evolution of each edge of abnormal p53 pathway with the controller.

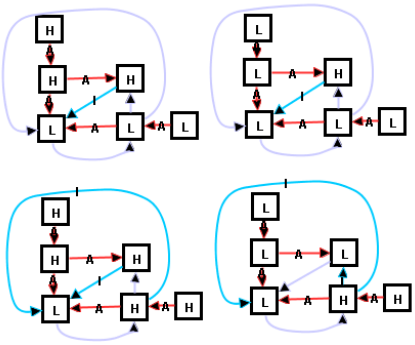


Figure 10: Possible cases for a normal p53 signal pathway where each box represents [ERK], [p19 ARF], [MDM2] and [P53] with different combinations of both [Ras] and [L26] where H represents an active state and L represents a non-active state [15].

V. HER2 Overexpressed Breast Cancer

We are interested in HER2 over-expressed breast cancer, which represents 20-30% of breast cancers. The experimental studies were done for investigating the effects of Tyrosine Kinase inhibitors (TKIs) on the BT474 and SKBr3 cell lines [16]. In this work, short term effects and long term effects of applying Gefitinib (a TKI) to those cell lines were studied and important effects of how the cancer cells overcome or escape from the inhibitory effects of TKIs were discovered. The authors in [16] showed that HER3 is recruited from the cytoplasm to the cell membrane to increase the triggering signal by vesicular trafficking in order to escape from HER2 inhibition. Also, they tested the effects of vesicular trafficking: when vesicular trafficking was stopped, phospho-HER3 and phospho-Akt did not survive from the inhibition of HER2.

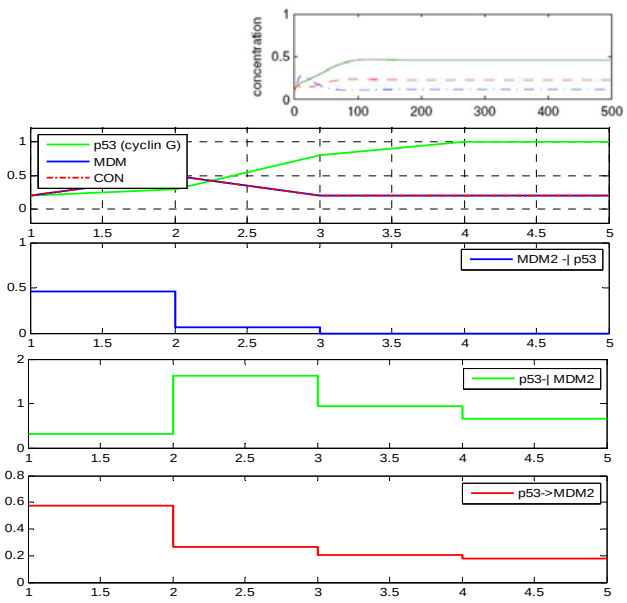


Figure 11: (upper) time course plots for the normal p53 pathway in Figure 4 (a) [15] (lower) the dynamic evolution of each edge of the normal p53 pathway.

We suspect there are short-term and long-term topological changes because TKI can inhibit and regulate downstream effectively in the short-term but it cannot regulate for the long-term. During the short-term, there is Positive Negative (PN) Feedback [17] so TKI inhibits HER3 effectively. However, for long-term behavior, even small triggering signal can amplify the phospho-Akt signal because of Positive Positive (PP) Feedback which is similar to vesicular trafficking. On the other hand, if the topology does not change, TKI should be able to regulate downstream over the long-term even though HER3 is recruited by vesicular trafficking.

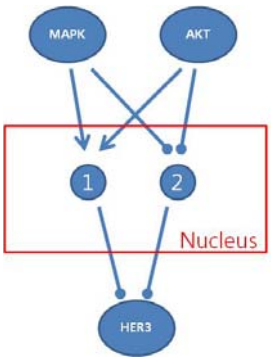


Figure 12: Bifan motif of nucleus, which is two-layered graphs with edges from nodes in top- to bottom-layer

We define the *a priori* map from biological information [16][18][19] where we include a 'nucleus' model to capture this possible topology change. The behaviors of the nucleus are not yet understood, however we abstract it with the switch as shown in Figure 12. Basically, there is a fail-safe mechanism, HER2-HER3 signaling which is buffered so that it is protected against an inhibition of HER2 catalytic activity and it is driven by the negative regulation of HER3 by Akt [18]. Also, there is a compensatory mechanism by cross-talk between MAPK and Akt which results in robust activation of this buffering. However, the compensatory buffering prevents apoptotic tumor cell death from occurring as a result of the combined loss of MAPK and Akt signaling [18]. For example, once a signal is triggered and either MAPK or Akt is high, then the nucleus stays "active" so MAPK and/or Akt are trying to keep the compensatory buffering. However, once both MAPK and Akt are down regulated, the nucleus is "deactivated" for all the time.

We apply the proposed optimization technique and the result is shown in Figures 13 and 14. Here, we use the generated data based on biological experimental data (western blot [16][18]). There are three main steps: before TKI is introduced (triggering network), right after TKI is introduced (short-term) and long term behavior after TKI is introduced. We can capture the topology change of biological network: for the initial stage (Figure 14 (a)), the signal is triggered and propagated along activation edges. After TKI is introduced, downstream components such as pHER3, pPI3K, Akt and MAPK are regulated because TKI inhibits and regulates downstream components. Moreover, the biological network shows PN Feedback which effectively modulates signal responses. Finally, for long term behavior, even if a small triggering signal is introduced (because of TKI inhibition, step 17-step 20), the downstream components are not regulated but are activated because the biological network evolves to Positive Positive (PP) Feedback which induces a slower but amplified signal response and enhances bi-stability.

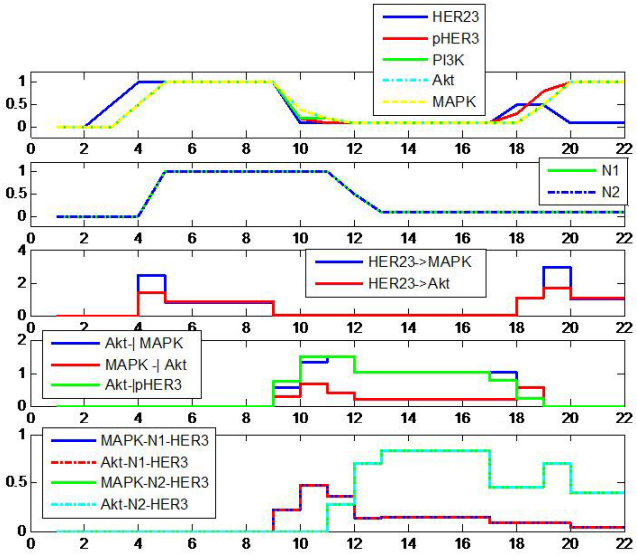


Figure 13: The upper two figures show the normalized biological data and the assumed nucleus level. The other (lower) figures show the strength of the downstream edges. For example, the edge connecting HER23 to MAPK (middle figure) is activated from step 4 to step 9 but deactivated from step 9 to 18.

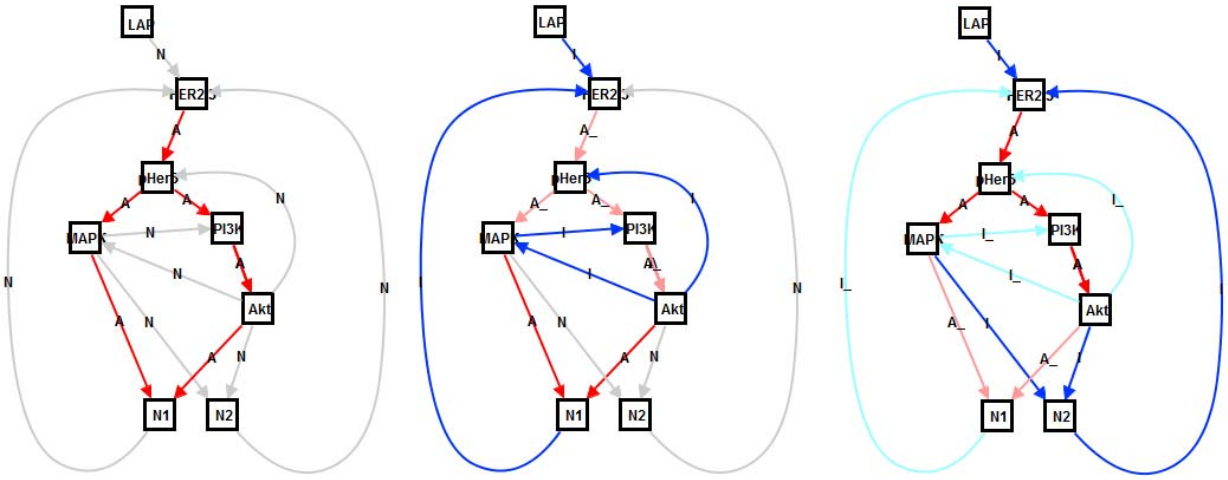


Figure 14: (a) Signal is triggered (b) TKI is introduced (short term) (c) TKI is introduced (long term)
(gray: not triggered edge, red : activation edge, blue : inhibition edge, light red/blue: deactivated edges after once activated)

VI. Conclusion

In this paper, we have proposed a convex optimization-based inference method in order to understand and identify a model for time evolving biological networks. The identification problem has led to a convex optimization problem with two main penalty functions by which we can match the experimental data with a sparse and robust representation, using *a priori* information of structure. We show that the proposed schemes can be useful to capture the dynamic evolution of the network and understand the biological system with a systems point of view, through examples. Also, we use this algorithm to study a breast cancer signal pathway to understand short-term and long-term behaviors.

Acknowledgment

The first author would like to thank support from STX foundation from Korea. This research was supported by the NIH NCI through the ICBP and PS-OC projects.

References

[1] K. Sachs, D. Gifford, T. Jaakkola, P. Sorger, and D. A. Lauffenburger. "Bayesian Network Approach to Cell Signaling Pathway Modeling", *Sci. STKE*, Vol. 2002, Issue 148, p. pe38, 3 September 2002

[2] N. Friedman and D. Koller. "Being Bayesian About Network Structure. A Bayesian Approach to Structure Discovery in Bayesian Networks", *MACHINE LEARNING*, Vol. 50, Numbers 1-2, pp 95-125

[3] J. Yu, V. A. Smith, P. P. Wang, A. J. Hartemink, and E. D. Jarvis. "Advances to Bayesian network inference for generating causal networks from observational biological data", *Bioinformatics* (2004) Vol. 20 (18), 3594-3603, July 29, 2004

[4] H. Schmidt, K.H. Cho, and E. W. Jacobsen. "Identification of small scale biochemical networks based on general type system perturbations", *FEBS J*. 272, Vol. 272, Issue 9, pp 2141–2151, May 2005

[5] M. Bansal, G. D. Gatta, and D. di Bernardo. "Inference of gene regulatory networks and compound mode of action from time course gene expression profiles", *Bioinformatics* (2006) Vol. 22 (7), pp 815-822

[6] H. Schmidt and E.W. Jacobsen. "Linear systems approach to analysis of complex dynamic behaviors in biochemical networks", *Systems Biology, IEE Proceeding*, Vol. 1, Issue 1, pp 149-158, June 2004.

[7] E. Sontag, A. Veliz-Cuba, R. Laubenbacher, and A. S. Jarrahy. "The effect of negative feedback loops on the dynamics of boolean networks", *Biophysical Journal*, Vol 95 (2), pp 518-526, July 2008

[8] Y. M. Zou, "Modeling and analyzing complex biological networks incooperating experimental information on both network topology and stable states", *Bioinformatics*, Vol 26 (16), pp 2037-2041, July 2, 2010

[9] U. Alon. "Network motifs: theory and experimental approaches", *Nature Review Genetics*, Vol 8, pp 450-461

[10] S. Han, Y. Yoon, and K. H. Cho, "Inferring biomolecular interaction networks based on convex optimization", *Computational Biology and Chemistry*, Vol 31, pp 347-354

[11] H. Y. Chang, B. H. Lawrence, and J. C. Diane. "Learning Patterns in the Dynamics of Biological Networks, International Conference on Knowledge Discovery and Data Mining archive", *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2009, pp 977-986

[12] M. S. Kim, J. R. Kim and K. H. Cho. " Dynamic network rewiring determines temporal regulatory functions in Drosophila melanogaster development processes", *Bioessays*, Vol 32, pp 505-513

[13] B.B. Marc, B. Alfredo, P. Andrea, W. Martin and Z. Riccardo. "Inference of sparse combinatorial-control networks from gene-expression data: a message passing approach", *BMC Bioinformatics* 2010, Vol 11, pp 355

[14] S. Boyd and L. Vandenberghe. 2004. Convex Optimization (Cambridge University Press)

[15] A. Aswani, N. Boyd, and C. Tomlin, "Graph-Theoretic Topological Control of Biological Genetic Networks", *American Control Conference* St. Louis, MO, USA June 10-12, 2009

[16] N. V. Sergina, M. Rausch, D. Wang, J. Blair, H. Byron, K. M. Shokat, and M. M. Moasser. "Escape from HER family tyrosine kinase inhibitor therapy by the kinase-inactive HER3", *Nature* Vol 445, 2007

[17] D. Kim, Y.K. Kwon, and K.H. Cho, "Coupled positive and negative feedback circuits form an essential building block of cellular signaling pathways", *BioEssays*, Vol 29, Issue 1, pp 85–90, January 2007

[18] D. N. Amin, N. Sergina, D. Ahuja, M. McMahon, J. A. Blair, D. Wang, B. Hann, K. M. Koch, K. M. Shokat, and M. M. Moasser, "Resiliency and Vulnerability in the HER2-Her3 Tumorigenic Driver". *Sci Transl Med* Vol 2, Issue 16, 2010

[19] S. Itani, J. Gray, and C. Tomlin, "An ODE Model for the HER2/3-AKT Signaling pathway in cancers that overexpress HER2", *American Control Conference*, St. Louis, MO, USA, June 10-12, 2009