Generating Stochastic Gene Regulatory Networks Consistent With Pathway Information and Steady-State Behavior

Jason M. Knight*, *Student Member, IEEE*, Aniruddha Datta, *Fellow, IEEE*, and Edward R. Dougherty, *Fellow, IEEE*

***Abstract*—We present a procedure to generate a stochastic genetic regulatory network model consistent with pathway information. Using the stochastic dynamics of Markov chains, we produce a model constrained by the prior knowledge despite the sometimes incomplete, time independent, and often conflicting nature of these pathways. We apply the Markov theory to study the model's long run behavior and introduce a biologically important transformation to aid in comparison with real biological outcome prediction in the steady-state domain. Our technique produces biologically faithful models without the need for rate kinetics, detailed timing information, or complex inference procedures. To demonstrate the method, we produce a model using 28 pathways from the biological literature pertaining to the transcription factor family nuclear factor-κB. Predictions from this model in the steady-state domain are then validated against nine mice knockout experiments.**

***Index Terms*—Gene regulatory networks, pathways, systems biology, stochastic modeling.**

I. Introduction

BIOLOGICAL regulatory network models offer the promise of applying system-based approaches for cancer diagnosis and therapy [1], [2]. With the advent of high-throughput genomic and proteomic data a large number of data-driven inference procedures have been proposed to infer regulatory networks and there are a number of reviews available in the literature [3]–[7]. Inference is inherently difficult owing to the combination of model complexity and limited data. Before the advent of high-throughput measurement techniques such as microarrays, biological experimentation often focused on uncovering (mostly univariate) relationships between genes and proteins in the production of pathway knowledge. This pathway knowledge is based on empirical observations across different experiments that have acquired some degree of validity through the peer review process. While not all pathways in the literature are accurate, and some pathways may in fact be conflicting, we believe that they offer an excellent foundation for network construction. Pathway information can reduce the data requirement and generate models that produce more reliable predictions than those with little or no prior information. Indeed, beginning with pathway knowledge, one can build networks based solely on that knowledge, thereby arriving at an uncertainty class of networks consistent (in some defined way) with the prior knowledge [8]. In this paper, we will extend the approach taken in [8] to produce a single stochastic network capturing the behavior of pathways. To illustrate the methodology, we will build a network related to the NF-κB system and, in a manner akin to that taken by [9], will validate the network by comparing its long-run behavior with known long-run behavior in the literature.

In [8], Layek *et al*. propose a procedure to infer deterministic dynamical models from Boolean pathway knowledge, which resolves pathway inconsistencies by relaxing pathway timings. In the case of incomplete knowledge, the procedure outputs an uncertainty class of deterministic models. In this paper, we propose that inconsistencies and incompleteness be incorporated into a single stochastic dynamical model of the system. Our model handles a variety of underlying pathway inconsistencies stemming from timing overlaps, different cellular contexts, and incomplete knowledge regarding pathways. By producing a single stochastic model rather than an uncertainty class of models, we can perform steady-state analysis and validation in a more straightforward way.

Some key assumptions underlie our model development. The most important assumption is the discrete-state and discrete-time approximation of protein behavior. It has been validated in many biological contexts [10], [11] and provides an important simplification that enables network modeling with pathway data. Indeed, such discrete-time, discrete-state modeling avoids the need for making continuous-time measurements of protein concentrations and facilitates the accommodation of genes/proteins which exhibit ON/OFF switch-like behavior. Moreover, discrete time systems are easier to analyze, model, and control in real time [12].

The second assumption is that there is no prior knowledge about the initial state of the cell; only the presence or absence of external stimuli is known. Consequently, we assign a uniform

Manuscript recieved November 11, 2011; revised Febraury 27, 2012; accepted March 18, 2012. Date of publication April 3, 2012; date of current version May 18, 2012. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant # A-7739, in part by the National Science Foundation under Grant ECCS-0701531, and in part by the W.M. Keck Foundation. *Asterisk indicates corresponding author*.

*J. M. Knight is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: jknight@tamu.edu).

A. Datta is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA. (e-mail: datta@ece.tamu.edu).

E. R. Dougherty is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA, the Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ 85004 USA, and the Department of Bioinformatics and Computational Biology of the University of Texas M. D. Anderson Cancer Center, Houston, TX 77030 USA (e-mail: edward@ece.tamu.edu).

Digital Object Identifier 10.1109/TBME.2012.2192117

initial distribution to all possible states and allow the pathway constrained model to stabilize into attractor cycles. This assumption has the effect of distributing the resulting probability mass of the cell states to states in the attractor cycles. Moreover, each attractor cycle has a total probability proportional to the size of its basin of attraction. A similar conclusion was reached by [13] but from a biological argument of cellular homeostasis in the presence of continual perturbations of the inter- and intracellular environments.

Overall, in this paper, we propose a new method for generating Markov chain network models from pathway information and apply this method to the set of proteins composing the NF-κB regulatory network to build the transition probability matrix of a discrete-state, discrete-time Markov chain that produces predictions that agree with the literature. The NF-κB system has been chosen due to the prevalence of associated pathway information in the literature as well as due to its biological importance in cancer and the innate immune system.

II. Methods

In network inference, the first step is selecting a specific biological system and choosing the specific agents for inclusion in the model. In this paper, this task is performed manually, although future work could include using selection techniques such as statistical tests on high-throughput data to identify the most relevant molecules for state-based modeling. The model generated in this paper consists of protein species with the exception of one lipoglycan [lipopolysaccharide (LPS)]. In general, the species in the pathways and the resulting model can be a mixture of any types of biological entities as long as the pathways accurately reflect the relationships being studied. Throughout, we will refer to the elements in the model as proteins with the tacit understanding that sometimes other biological entities would also be admissible.

A. Pathways

To obtain pathways, we manually reviewed the biological literature relevant to the NF-κB system and recorded a pathway when significant biological evidence was available to indicate that the molecules involved in the pathway were significant. For interactions that include nonsignificant species, the pathway was either ignored or extended upstream and downstream until it included significant species. The resulting pathways and the references used to arrive at them are summarized in Table I. Each pathway description consists of two parts, the predicate and the subject separated by the implication sign, $\Longrightarrow$. The information that the pathway contains can be understood as: "when the predicate is true, the subject is implied to occur in the future." The timing with which this dependence occurs is not known. We assume that the dependence relationship is implemented at the next time step as in [8]. A full description of the NF-κB system appears later in Section II-F. Using the pathway data, one can determine which proteins are upstream of each other and also determine the set of predictor proteins, i.e., the proteins whose activity status collectively determines the time course updates of a given protein.

TABLE I
NF-κB System Pathways

Pathway	Reference
RIP1 $= 1 \Longrightarrow$ NEMO $= 1$	[15]
A20 $= 1 \Longrightarrow$ NEMO $= 0$	[15]
LTβR $= 1 \Longrightarrow$ NEMO $= 1$	[15]
A20 $= 1$ and RIP1 $= 1 \Longrightarrow$ NEMO $= 0$	[15]
RIP1 $= 0$ and LTβR $= 0 \Longrightarrow$ NEMO $= 0$	[15]
TNFR $= 1 \Longrightarrow$ AP-1 $= 1$	[15]
TNFR $= 0 \Longrightarrow$ AP-1 $= 0$	[15]
TNFR $= 1 \Longrightarrow$ RIP1 $= 1$	[18]
TNFR $= 0 \Longrightarrow$ RIP1 $= 0$	[18]
LPS $= 1 \Longrightarrow$ TNFR $= 1$	[19]
TNF$\alpha = 1 \Longrightarrow$ TNFR $= 1$	[19]
LPS $= 0$ and TNF$\alpha = 0 \Longrightarrow$ TNFR $= 0$	[19]
IKK$\alpha = 1 \Longrightarrow$ p52 $= 1$	[15]
IKK$\alpha = 0 \Longrightarrow$ p52 $= 0$	[15]
LTβR $= 1 \Longrightarrow$ IKK$\alpha = 1$	[15]
NEMO $= 1 \Longrightarrow$ IKK$\alpha = 1$	[15]
NEMO $= 0$ and LTβR $= 0 \Longrightarrow$ IKK$\alpha = 0$	[15]
NEMO $= 1 \Longrightarrow$ IKK$\beta = 1$	[20]
LPS $= 1 \Longrightarrow$ IKK$\beta = 1$	[18]
NEMO $= 0$ and LPS $= 0 \Longrightarrow$ IKK$\beta = 0$	[18], [20]
p65 $= 1 \Longrightarrow$ IκB $= 1$	[15], [21]
p65 $= 1 \Longrightarrow$ A20 $= 1$	[15], [21]
p65 $= 0 \Longrightarrow$ A20 $= 0$	[15], [21]
IKK$\beta = 0 \Longrightarrow$ IκB $= 1$	[15]
IKK$\beta = 1 \Longrightarrow$ IκB $= 1$	[15]
IκB $= 1 \Longrightarrow$ p65 $= 0$	[15]
IκB $= 0 \Longrightarrow$ p65 $= 1$	[15]
IKK$\alpha = 1 \Longrightarrow$ p65 $= 0$	[22]

B. Small Example

We now demonstrate this algorithm on a part of the NF-κB system. From Table I in Section II-F, the first five entries describe the behavior of the NF-κB essential modulator (NEMO) protein and are given by

$$\text{RIP1} = 1 \Longrightarrow \text{NEMO} = 1 \tag{1}$$

$$\text{A20} = 1 \Longrightarrow \text{NEMO} = 0 \tag{2}$$

$$\text{LT}\beta\text{R} = 1 \Longrightarrow \text{NEMO} = 1 \tag{3}$$

$$\text{A20} = 1 \text{ and } \text{RIP1} = 1 \Longrightarrow \text{NEMO} = 0 \tag{4}$$

$$\text{RIP1} = 0 \text{ and } \text{LT}\beta\text{R} = 0 \Longrightarrow \text{NEMO} = 0. \tag{5}$$

These mandate the following relationships: when receptor-interacting protein 1 (RIP1) is activated, it activates NEMO; when A20 is activated, it deactivates NEMO; when LTβR is activated, it activates NEMO; when both A20 and RIP1 are activated, NEMO is deactivated; and when RIP1 and LTβR are both inactive, NEMO is deactivated. From these, we infer that a reasonable predictor set for NEMO is $\{\text{A20}, \text{LT}\beta\text{R}, \text{RIP1}\}$. One can similarly find predictor sets for the other biological entities in Table I.

Having identified the predictor set for each protein, we can use a Karnaugh map [14] to determine the update rule for that protein. The method used is an extension of the one developed in [8]. For each protein, using its predictor set, we initialize a Karnaugh map with every entry in the map containing the unknown "x." This is consistent with the fact that to start we lack information on how the current values of the prediction proteins affect the update value of the predicted protein at the next time step. The Karnaugh map entries can then be updated

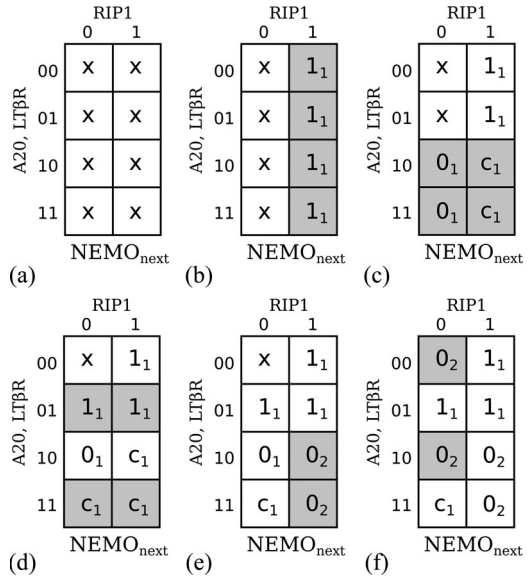


Fig. 1. Process of developing the Karnaugh map using the pathways associated with NEMO. For each step in the process, we evaluate each pathway in turn. In each pathway, the predicate specifies certain locations in the Karnaugh map and these are shaded in gray in the corresponding table.

by incorporating the pathway information. For instance, consider the update of NEMO. The initial blank Karnaugh map for NEMO is shown in Fig. 1(a). We refer to the locations in the Karnaugh map using the vector [A20, LTβR, RIP1]. Thus, 000 corresponds to the square in the upper left corner of the map and so on. Using the first pathway, RIP1 $= 1 \implies$ NEMO $= 1$, we can fill all the entries that correspond to RIP1 = 1 with 1_1. This results in the Karnaugh map shown in Fig. 1(b). The subscript 1 indicates that the conflict comes from a pathway containing only one protein in the predicate. The reasoning for this notation will soon become clear and is related to the specificity of the pathway.

We next proceed to fill the locations 100 and 110 with 0_1 to satisfy the requirements of the second pathway, A20 $= 1 \implies$ NEMO $= 0$. Here, we face a conflict when RIP1 $= 1$ and A20 $= 1$. At the two conflicting locations, we replace the 1s with c_1. The letter c indicates a conflict at this location. This results in the table in Fig. 1(c).

Now consider pathway three, LTBR $= 1 \implies$ NEMO $= 1$, which mandates that the x at location 010 be replaced by 1_1. Since location 110 contains a 0_1 and the subscript of the currently considered pathway is no larger than the present value, we replace it with c_1 to indicate a conflict. Finally, location 111 already has c_1 and, because this pathway only contains one protein in its predicate, we leave the c_1 in place. This leads to the table shown in Fig. 1(d).

Now consider the fourth pathway, A20 $= 1$ and RIP1 $= 1 \implies$ NEMO $= 0$. The predicate here applies to locations 101 and 111. Both of these contain c_1 conflicts, but because this pathway contains two proteins in its predicate, we acknowledge this pathway has more specific information regarding this particular experimental scenario and can override the c_1 conflicts. Therefore, we fill the locations 101 and 111 with 0_2 and obtain the table shown in Fig. 1(e).

For the final pathway, RIP1 $= 0$ and LTBR $= 0 \implies$ NEMO $= 0$, the predicate applies to the locations 000 and 100. The former has an x, so we replace it with 0_2, and the latter contains 0_1, which is of a lower specificity than this pathway, so we replace it with 0_2. NEMO is now finished and have the Karnaugh map of Fig. 1(f), in which there is only one uncertainty condition at the location 110.

C. General Procedure

The procedure demonstrated on the NEMO example can be generalized by proceeding through the pathways in order from the least specific to the most specific predicates and filling in the Karnaugh map entries if information is available or invalidating information already provided if there are conflicts in the pathway information. In the NEMO example there were no x values remaining at the end of the procedure. Given less information there could be x's remaining in the Karnaugh map and these would be treated in the same manner as a remaining conflict: a random selection is made. This general procedure is presented in Fig. 2. The key difference between the algorithm presented in [8] and the one here is that in [8], one attempts to resolve the conflicting entries in the Karnaugh maps by suitably altering the timings of some of the pathways, whereas here the conflicts in the Karnaugh maps are retained. We take the view that there is no sure way of resolving conflicts via timing alteration based on the prior knowledge contained in the pathways. Hence, the state transitions following a conflict will not be unique and, assuming that all the subsequent states are equiprobable, we derive a probabilistic state transition graph (PSTG) for the system.

A PSTG is a directed graph describing the evolution of a system through time. It consists of k^n nodes corresponding to the states of the system, where k is the number of quantization levels associated with the activity state of each protein (assumed throughout to be two for a binary quantization) and n is the number of proteins in the system. Additionally, each directed edge indicates a viable transition between states as allowed by the pathway information.

For example, using the NEMO pathways and the Karnaugh map generated in the last section we obtain a simplified but illustrative example of a PSTG for this system as shown in Fig. 3. We assume here that the predictor proteins have no predictors themselves and, therefore, exhibit static behavior. This is seen in Fig. 3 where the edges of the PSTG, or the allowed transitions, are between states with identical values for the predictor proteins and only differing in the values of NEMO. Also, based on the Karnaugh map generated from the NEMO pathways, we expect uncertainty at the state $110\times$ and indeed, both the states 1100 and 1101 have two outgoing edges each. One of these creates a self loop and the other directs to the corresponding state with the value of NEMO flipped.

The PSTG can be viewed as a compact representation for the class of networks which the uncertain Karnaugh maps such as the one in Fig. 1(f) generate. To use the preceding example, we could also represent the uncertainty of NEMO at $110\times$ by

Require: x: protein to produce a Karnaugh map for
Require: P: sorted list of pathway segments as described in [8] in ascending number of predictive proteins
1: **for all** p $\leftarrow$ P **do**
2: $S \leftarrow$ Subjects(p)
3: **if** $x \in S$ **then**
4: PredictorSet $\leftarrow$ PredictorSet $\cup$ PredicateProteins(p)
5: **end if**
6: **end for**
7: **for all** p $\leftarrow$ P **do**
8: $S \leftarrow$ Subjects(p)
9: **if** $x \in S$ **then**
10: $n \leftarrow$ |PredicateProteins(p)|
11: $C \leftarrow$ PredicateCondition(p)
12: $v \leftarrow$ NextStateValue(p)
13: **for all** value permutations e of PredictorSet **do**
14: **if** $(K(e) = \mathrm{c}_z$ **or** $K(e) = \mathrm{x}$ **or** $K(e) = y_z)$ **and** $C \subset e$ **then**
15: K(e) $\leftarrow v_n$
16: **else**
17: K(e) $\leftarrow \mathrm{c}_n$
18: **end if**
19: **end for**
20: **end if**
21: **end for**
22: **return** K, the resulting Karnaugh map

Fig. 2. This algorithm produces a Karnaugh map for a given protein from a sorted list of pathways. It is simple and optimized for clarity of exposition rather than run time. Therefore, we use two loops through the pathways. The first is to collect the set of predictor proteins for this protein, and the second is to fill in the entries of the Karnaugh map. Before the second loop we must initialize the Karnaugh map with x's in each location. On lines 10–12 we break up the pathway into its constituent parts, and the loop starting on line 13 iterates over the possible permutations of the values of the predictor set for this protein. Beginning on line 14, if the permutation e matches the condition of the predicate for this pathway, and if this pathway can override the information already in the Karnaugh map (as it is more specific with $n > z$), then overwrite it with the next state value subscripted with the specificity of this pathway. Otherwise mark the location as a conflict with this pathway's specificity.

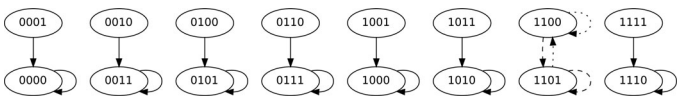


Fig. 3. Simplified PSTG for the NEMO protein with predictors that are all static for the purposes of illustration. Here the binary value of the state should be interpreted as [A20, LTβR, RIP1, NEMO]. The two dashed edges represent one possible configuration of the network, while the two dotted edges represent the other configuration.

creating two separate networks with two corresponding state transition diagrams. One network would contain the two dotted edges in Fig. 3 and predict that NEMO should equal 0 when A20 = 1, LTβR = 1, and RIP1 = 0, while the other would have the two dashed edges and predict that NEMO should equal one for the same set of predictor values.

To demonstrate the process of converting Karnaugh maps to PSTGs, we use the synthetic example in Fig. 4(a), where we are given five pathways for the three proteins A, B, and C. These produce the two Karnaugh maps in Fig. 4(b) and the resulting

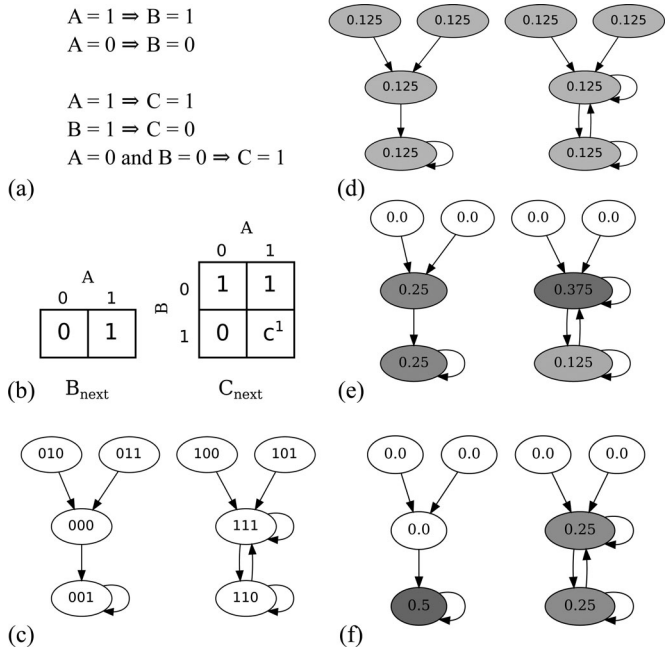


Fig. 4. Synthetic example of the complete process from pathways to long run probabilities. (a) Pathways for the three proteins: protein A has no predictors, B is predicted by A, and C is predicted by A and B. (b) Resulting Karnaugh maps with one final conflict obtained using the method described in Fig. 2. (c) Resulting PSTG shows two separate basins. The left basin has a single attractor state 001 while the right basin has two states in the attractor, 111 and 110. (d)–(f) Probability mass in each state at the (d) initial stage, (e) after the first state transition, and (f) after many state transitions have taken place.

PSTG is shown in Fig. 4(c). Protein *A* has no predictors and thus the state space can be partitioned into two basins according to its activity.

To produce the PSTG in Fig. 4(c), we begin by listing all the possible 2^3 states as unconnected nodes in a graph. Then, for each node we use the Karnaugh maps to update the status of each gene and then concatenate this information to determine the next state to which the model will transition. For example, for the state 000, protein A has no predictors and, therefore, no Karnaugh map, so we assume it remains in the state 0. Using the Karnaugh map for B, A is currently 0 so that $\mathrm{B}_{\mathrm{next}} = 0$. Finally, proteins A and B being zero implies that $\mathrm{C}_{\mathrm{next}} = 1$. Combining all this information creates a transition edge from state 000 to 001 in the PSTG.

As another example, consider the state 111. Following the aforementioned logic, $\mathrm{A}_{\mathrm{next}} = 1$ and $\mathrm{B}_{\mathrm{next}} = 1$, but $\mathrm{C}_{\mathrm{next}}$ is a conflict, so 111 will progress to $11\times$, which can be expanded to 110 and 111. This is shown in Fig. 4(c), where the node 111 has two outgoing edges, a self loop and an edge linking to 110.

The PSTG, as in Fig. 4(c), can also be interpreted as the state transition graph of a Markov chain. Thus, the Karnaugh maps can be converted into a $2^n \times 2^n$ ($n = 3$ here) transition probability matrix $\mathbf{Q} = [p_{i,j}]_{2^n \times 2^n}$, where each entry $p_{i,j} = P(X_{n+1} = j | X_n = i)$ represents the probability of the model transitioning from state i to state j in one time step. This matrix is stochastic and is a row normalized sum of the individual deterministic state transition matrices of the resulting class of Boolean networks that would be required to accommodate the

uncertainties associated with the different Karnaugh maps. We will use this state transition matrix to simulate the long run behavior of the model.

D. Long-Run Behavior

A salient objective of the model is to obtain a useful measure of the long-run behavior of the system akin to the steady-state distribution for an ergodic Markov chain. Since the PSTG does not necessarily provide an irreducible or aperiodic state space, we propose a method to approximate its long-run behavior.

To produce the state transition matrix $\mathbf{Q}$, we postulate that any state having more than one outgoing edge will have a transition probability evenly split among the possible transitions. Thus for state 110 in the preceding example we have $p_{6,6} = P(X_{n+1} = 6|X_n = 6) = 0.5$ and $p_{6,7} = P(X_{n+1} = 6|X_n = 7) = 0.5$, with all other entries in the sixth row of $\mathbf{Q}$ being zero. For a deterministic state such as 100, all entries in the four row of $\mathbf{Q}$ are zero except for $p_{4,7} = P(X_{n+1} = 7|X_n = 4) = 1.0$.

In general, we do not know the initial state of the system, so we assume uniform probabilities for the initial state vector x_0. For example, in Fig. 4(d), the initial state is

$$x_0 = [0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125].$$

Now, by the properties of Markov chains, we can transition the system by left multiplying

$$x_1 = x_0\mathbf{Q} = [0.25, 0.25, 0, 0, 0, 0, 0.125, 0.375]$$

which gives Fig. 4(e). Via matrix exponentiation to see that

$$x_2 = x_1\mathbf{Q} = x_0\mathbf{Q}^2 = [0, 0.5, 0, 0, 0, 0, 0.25, 0.25]$$

as seen in Fig. 4(f). The system has settled to an invariant measure as $x_2 = x_2\mathbf{Q}^n \ \forall n \in \mathbb{N}^+$.

In general, repeating this matrix exponentiation will result in the probability mass accumulating in the attractors of the system, as in the example; however, in the case of systems with periodic attractors, the masses will not converge to a limiting probability mass. Instead, the probability mass might oscillate between states or around a cycle indefinitely, analogous to oscillatory behavior in undamped systems. To overcome this problem, it is necessary to first propagate the mass a sufficient number of times to eliminate all the transient states and then average the probability masses from the set of tail states

$$\pi \approx \frac{1}{n}\sum_{k=i}^{i+n} x_0\mathbf{Q}^k \tag{6}$$

where i is sufficiently large to move the probability mass into the attractors, and n is sufficiently large to smooth out any oscillations in the probability mass distribution in the attractor cycles.

For demonstration, assume that a set of pathways produces a Markov chain transition graph as shown in Fig. 5(a). The state transition matrix for this Markov chain is given by

$$\mathbf{Q} = \begin{pmatrix} p_{00,00} & p_{00,01} & p_{00,10} & p_{00,11} \\ p_{01,00} & p_{01,01} & p_{01,10} & p_{01,11} \\ p_{10,00} & p_{10,01} & p_{10,10} & p_{10,11} \\ p_{11,00} & p_{11,01} & p_{11,10} & p_{11,11} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

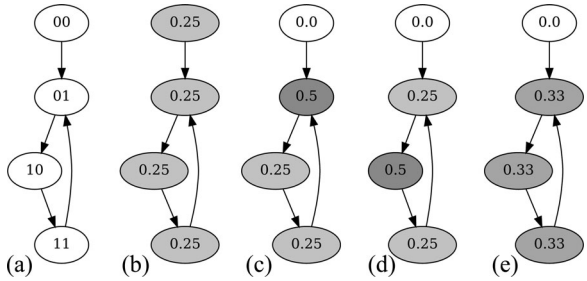


Fig. 5. State transition graph of a Markov chain and the simulation of its long run probabilities. (a) Graph and the names of its nodes. (b)–(d) Simulation of the Markov chain after its initialization, first, and second step, respectively. (e) Long run probability limit after sufficient averaging.

where $p_{00,01} = P(X_{n+1} = 01|X_n = 00)$. Beginning with a uniform probability distribution, we initialize $x_0 = [0.25, 0.25, 0.25, 0.25]$ and simulate the progression of time by evaluating the vector matrix products

$$x_0 = [0.25, 0.25, 0.25, 0.25]$$
$$x_1 = x_0\mathbf{Q} = [0.0, 0.5, 0.25, 0.25]$$
$$x_2 = x_1\mathbf{Q} = [0.0, 0.25, 0.5, 0.25]$$
$$x_3 = x_2\mathbf{Q} = [0.0, 0.25, 0.25, 0.5]$$
$$x_4 = x_3\mathbf{Q} = [0.0, 0.5, 0.25, 0.25].$$

We can see that x_0 is visualized in Fig. 5(b), x_1 is visualized in Fig. 5(c), and x_2 is visualized in Fig. 5(d). We also see that $x_4 = x_1$ and that this system is, therefore, periodic. This means that the limit $x_\infty = \lim_{n\to\infty} x_0\mathbf{Q}^n$ does not exist and we must perform a smoothing operation to approximate the long run behavior.

Following the smoothing summation in (6)

$$\pi = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{i+n} x_0\mathbf{Q}^k \approx [0.0, 0.33, 0.33, 0.33]$$

with $i = 1$ to eliminate the mass in the transient states. The result of this smoothing is shown in Fig. 5(e). With this approximation of long run probabilities, we will now transform this into a more biological relevant quantity.

E. Steady-State Activities

Biologically, it is more relevant to determine how often a particular protein is active instead of determining how much time is spent in a particular state. Accordingly, we apply a transformation to the long-run probability approximation and define a *steady-state activity (SSA) vector* with n components, indexed from 0 to $n-1$, corresponding to the n proteins in the model. The ith component of the SSA vector, characterizing the activity of the ith protein, is computed as

$$\text{SSA}(i) = \sum_{j=0}^{2^n-1} \pi_j z_i(j)$$

where $z_i(j)$ is the binary value of the ith protein in the jth decimal state, and π_j is the jth entry of the approximation to the

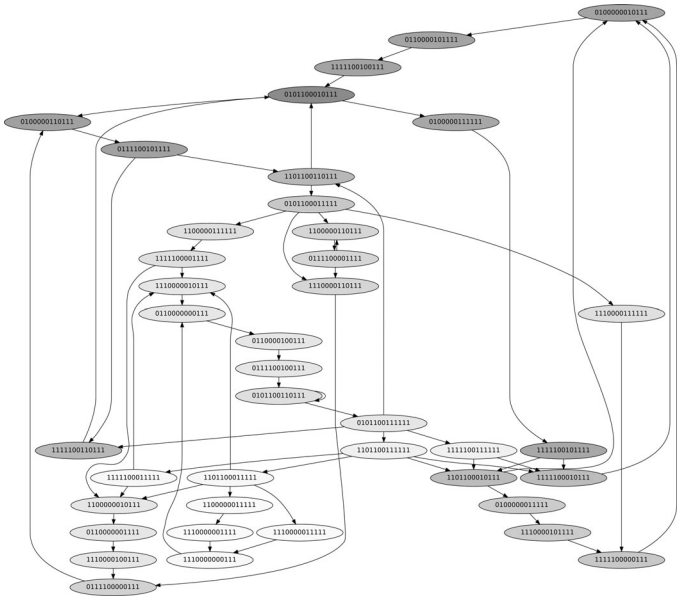


Fig. 6. Resulting communicating class of states for the full NF-κB PSTG when the stimuli conditions are set to TNF=1, LPS=0, and LTβR=0. The 13 bit binary vector can be read as [A20, AP-1, IκB, IKKα, IKKβ, LPS, LTβR, NEMO, p52, p65, RIP1, TNFα, TNFR]. To give a walk through of one state transition, starting at the upper right state, 0100000010111, many things occur in one transition: IκB is inactive and thus at the next state p65 translocates to the nucleus to become active; NEMO is activated by RIP1; p52 is deactivated as IKKα is not activating it; and IκB becomes active, as constitutive expression allows it to repopulate the cytoplasm in the absence of activated IKKβ . All of these changes results in the model evolving to state 0110000101111.

long-run behavior of the system just presented. $\mathrm{SSA}(i) \in [0, 1]$ and gives the probability that the model is likely to exist in states where protein i is active. For example, $\mathrm{SSA}(i) = 1$ indicates that the ith protein is active in every attractor state.

Applying this to our example in Fig. 4(f), the SSA for protein A, i.e., $\mathrm{SSA}(0)$, is calculated by considering the two attractor states 111 and 110 with active protein A as shown in the right-hand basin in the figure. Therefore, $\mathrm{SSA}(0) = 0.25 + 0.25 = 0.5$. Similarly, for protein B, $\mathrm{SSA}(1) = 0.5$, while for protein C, $\mathrm{SSA}(2) = 0.75$, the latter being due to the fact that all attractors have protein C active except the state 110, which has a final mass of 0.25.

To appreciate the impetus for using the SSA vector instead of the state vector, consider the network constructed by applying Fig. 2 to the set of 28 pathways involving NF-κB. With the external stimuli set to TNF$\alpha = 1$, LPS $= 0$, and LTβR $= 0$, only the communicating set of states for the resulting PSTG is displayed in Fig. 6. Due to the size and complexity of the resulting set of states, it would be difficult, if not impossible, to compare the behavior of this PSTG with that obtained from any of the knockout experiments. The SSA vector ameliorates this problem and aids in carrying out qualitative comparisons with the experimental data.

F. NF-κB System

Nuclear factor-κB (NF-κB) is a protein dimer from the rel family of transcription factors that promote the expression of over 100 genes, primarily in the immune system [16]. The NF-κ B system's primary role in the immune system is in the production of inflammatory cytokines, small signaling proteins used extensively in cell-to-cell communication. NF-κB also has both proapoptotic and antiapoptotic effects on the cell and the balance of these responses can be adjusted by stimuli.

The NF-κB transcription factor is a key element in the inflammation stress response pathway. The general architecture of this system is typical of several stress response pathways [17]. NF-κB is sequestered in the cytosol by the "sensor," which in this case is Iκ B, and when degraded by the "transducer" of IKKβ, it allows for a rapid downstream response without the lag associated with *de novo* protein synthesis. This combination of a transducer, sensor, and transcription factor is a common motif seen in stress response pathways and forms the backbone of the NF-κB system [17].

The mammalian NF-κB family consists of p65 (RelA), RelB, c-Rel, NF-κB1 (p52/p100), and NF-κB2 (p50/p105). NF-κB is constitutively expressed, but sequestered in the cytosol by a family of IκB inhibitor proteins which include the p100 and p105 precursors to p52 and p50, respectively, along with IκBα, IκBβ, IκBϵ, IκBγ, and BCL-3 [15]. These IκBs prevent the NF-κB dimers from reaching their binding sites in the nucleus.

The IκB kinase complex (IKK) consists of the IKKα, IKKβ, and IKKγ (NEMO) subunits. The NEMO subunit is a regulator and maintains the IKKα and IKKβ subunits in inactive states.

The signaling pathways involved in the NF-κB system are shown in Fig. 7 and listed in Table I. NF-κB activation occurs through two separate cascade pathways. The canonical pathway is primarily activated by the proinflammatory cytokine tumor necrosis factor-α (TNFα). When TNFα binds to TNF receptor protein (TNFR), it begins a signaling cascade that through the RIP1 activates the NEMO subunit of IKK, which activates both the IKKα and IKKβ subunits [18]. The IKKβ subunit then proceeds to phosphorylate I κB proteins, which leads to their destruction through polyubiquitination and allows NF-κB dimers, primarily p65 heterodimers and homodimers, to translocate to the nucleus and bind to promoter regions [15].

Bacterial LPS is a component of bacterial cell walls that provides an activating stimulus for Toll-like receptors 2 and 4 (TLR2 and TLR4). These receptors also activate the canonical pathway but through MyD88 and Trif intermediary proteins [19] that directly activate the IKKβ subunit without activating NEMO [18]. Also, the LPS-dependent pathway indirectly activates the canonical pathway through autocrine stimulation via the production of TNFα. The alternative pathway is activated through CD40 and LTβ R, and via the NIK protein directly activates the IKKα subunit, which via phosphorylation, processes p100 into p52, which activates the nuclear localization segment (NLS), thereby allowing the p52 dimer to translocate into the nucleus.

NF-κB activates two genes that produce IκB and A20. These act as negative feedback to dampen the response of the canonical pathway. A20 binds to NEMO and impairs the activation of the IKKβ and IKKα subunits by RIP1 [20]. Additionally, NF-κB activates antiapoptotic genes such as cFLIP (not shown in Fig. 7) that counteract the TNFα induced activation of the proapoptotic

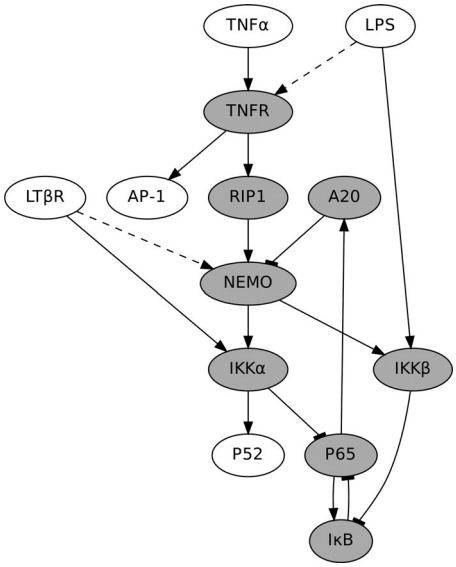


Fig. 7. Pathway structure of the NF-κB system. Gray shaded proteins are those that are knocked out in the validation section of this paper. The presence of a directed edge indicates that a pathway exists that shows the upstream protein causes a change in the activity of the downstream protein. Inhibitory pathways are terminated with a filled rectangle. The two dotted connections denote special interactions: The LPS induced autocrine production of TNFα would seemingly imply that an excitatory connection should be made between LPS and TNFα. However, because we want to exogenously control TNFα, LPS, and LTβR in our knockout simulations, we consider TNFα to be an exogenous stimulus, thereby allowing us to control that level in simulations without affecting the autocrine feedback loop of LPS. The second dotted connection is the pathway from LTβR to NEMO as it is a result of an unknown molecular mechanism but described in [15].

AP-1 family, such as c-Jun [21]. Thus, NF-κB knockouts are likely to exhibit apoptotic behavior when subjected to TNFα stimulation.

III. Results and Discussion

The model developed in this paper has been designed to preserve the biological state transitions and stable state attractor cycles. Thus, it is reasonable to validate our model using the experimentally observed long-run behavior of biological systems. Focusing on animal knockout models, we employ the genetic regulatory network model validation framework outlined in [22]. We use the direction of shift in SSA values after a protein is knocked out as the characteristic of our model with which to examine its validity. For animal model knockout experiments that measure protein species in the experimental cells directly, we compare the directions of these shifts to the shifts seen in our simulation. For knockout experiments that do not measure species concentrations, we need a different model characteristic to use for comparison. In these cases, we consider the qualitative shift in SSA values of the simulation and interpret what phenotype shift would result based on the known phenotypical characterizations of the individual proteins. This phenotype shift is then compared against the phenotype seen in the knockout experiment. The use of qualitative sign shifts as a characteristic can be seen in [9], where the possible transient sign changes the model could take were checked against the known biological sign transitions.

The goal of the inference procedure is not to produce novel pathways (that would need to be checked against "unseen" testing experiments) from existing ones, but to create a dynamic model that behaves in accordance with the combination of all the pathways. From a modeling perspective, we are assuming that the pathways constitute prior knowledge and no data are used in network construction. Thus, while we have aimed to minimize overlap in the studies from which the pathways have been taken and the studies in which the long-run effects of knockouts are observed, under the assumption that the pathways constitute prior knowledge, the purpose of the knockout observations is to validate the model constructed from existing (prior) knowledge by demonstrating the consistency of consequences of the network model with physical observations, which is a classical validation approach [23].

The PSTG resulting from the 28 NF-κB pathways in Table I was too large to be visually interpreted, although for illustrative purposes a small subset of it is included in Fig. 6. We have compared the behavior of our model in a steady-state fashion with the phenotypes and measured protein quantities as found in the knockout studies. Due to the qualitative nature of the data collected in the knockout experimental studies used, the comparisons between the model and the study will by necessity be qualitative. We believe that this still allows for preliminary validation (as in the case of [9] and as discussed in [22], given the complexity of the model, the inherently noisy nature of biological systems and experimentation, and the large number of knockout studies examined.

A. $A20^{-/-}$

Werner *et al*. [24] aimed to derive ordinary differential equation models of NF-κB regulation in response to TNFα and LPS stimulation. Their model includes the negative feedback of A20, and to justify this, they compared $A20^{+/+}$ against $A20^{-/-}$ Murine 3T3 immortalized fibroblasts and measured IKK and NF-κ B activity in response to 45 min of TNFα stimulation. It is clear in this comparison that the $A20^{-/-}$activity of IKK and NF-κ B is much higher than that of the $A20^{+/+}$ cells. This is consistent with Table II(a), produced by simulating our model, where the levels of IKKβ and p65 both increase when the model is constrained to $A20^{-/-}$. While the increase in p65 is small, the direction of the change is consistent with the findings of Werner *et al*.

B. $IKK\beta^{-/-}$ and $TNFR^{-/-}$

Li *et al*. [25] used $IKK\beta^{-/-}$ knockout mice to investigate the role of IKKβ in the NF-κB signaling pathway. They determined that the lack of IKKβ increases hepatocyte death due to TNFα (TNFα toxicity). They also found that the $IKK\beta^{-/-}$ knockout does not affect c-Jun levels, a member of the proapoptotic AP-1 family, which helps explain the increased toxicity. They also measured an increase in stability in IκB from $IKK\beta^{-/-}$ lines, which is seen in our model as an increase in the activity of IκB in Table II(b).

TABLE II
KNOCKOUT SIMULATION RESULTS UNDER TNFα AND LPS STIMULATION

Knockout	TNFα, LPS	Baseline SSA	Knockout SSA	Reference
(a) A20$^{-/-}$	1, 0	p65=0.39, p52=0.61, IKKβ=0.61	p65=0.40, p52=1.00, IKKβ=1.00	[28]
(b) IKK$\beta^{-/-}$	1, 0	p65=0.39, IκB=0.49, AP-1=1.00	p65=0.00, IκB=1.00, AP-1=1.00	[29]
(c) IKK$\beta^{-/-}$, TNFR$^{-/-}$	1, 0	p65=0.39, AP-1=1.00	p65=0.00, AP-1=0.00	[29]
(d) p65$^{-/-}$	1, 0	p65=0.39, AP-1=1.00	p65=0.00, AP-1=1.00	[25], [30]
(e) IKK$\alpha^{-/-}$	0, 1	p65=0.60, A20=0.60, IκB=0.30	p65=0.67, A20=0.67, IκB=0.33	[22], [31]
(f) IKK$\beta^{-/-}$	0, 1	p65=0.60, AP-1=1.00	p65=0.00, AP-1=1.00	[30], [32]
(g) NEMO$^{-/-}$ (macrophage)	0, 1	p65=0.60, AP-1=1.00	p65=0.67, AP-1=1.00	[33]
(h) NEMO$^{-/-}$ (general)	0, 1	p65=0.67, AP-1=1.00	p65=0.67, AP-1=1.00	[33]
(i) IκB$^{-/-}$	1, 0	p65=0.39, p52=0.61	p65=1.00, p52=0.00	[34], [35]
(j) IκB$^{-/-}$	0, 1	p65=0.60, p52=0.40	p65=1.00, p52=0.00	[35]
(k) RIP1$^{-/-}$	1, 0	p65=0.39, p52=0.61	p65=0.00, p52=0.00	[36]

They found that a IKK$\beta^{-/-}$, TNFR$^{-/-}$ double knockout, where both genes are knocked out simultaneously, allows the mice to survive to term (rescuing the phenotype). We mirror this result in our model as the IKK$\beta^{-/-}$, TNFR$^{-/-}$ double knockout in Table II(c), which shows the same reduction of p65 as the single knockout, but with the c-Jun (AP-1 family) activation also reduced, thereby reducing the pro-apoptotic nature of the TNFα stimulus under IKK$\beta^{-/-}$ knockout conditions. This explains the reduction in TNFα toxicity.

C. $p65^{-/-}$

Prendes *et al*. [21] used fetal liver hematopoietic precursors from mice embryos deficient in RelA (p65) to study the effect of RelA deficiency in lymphocytes. They found that the loss of RelA increased TNFα toxicity, which was ameliorated when cells were induced by virus to produce the antiapoptotic NF-κB target gene cFLIP, indicating that the increased cell death was due to the inhibition of NF-κB. This is backed by our knockout model by a reduction in RelA at the steady state in Table II(d).

D. $IKK\alpha^{-/-}$

Li *et al*. [26] used embryonic liver-derived macrophages (ELDM) from IKK$\alpha^{-/-}$ mice to determine the role of IKKα in the innate immune system's inflammation response. IKK$\alpha^{-/-}$ ELDM cells were found to exhibit higher than normal antigen presenting response and higher NF-κB levels in response to LPS stimulation. [27] used a model with an inactivatable variant of IKKα (denoted by IKK$\alpha^{A/A}$) and observed an increase in NF-κB and A20 upon the application of LPS, both of which match our model in Table II(e).

Also, Li *et al*. found a decrease in the postinduction response of IκB in their IKK$\alpha^{-/-}$ cells, whereas Lawrence *et al*. measured an increase in the amount of IκB for their IKK$\alpha^{A/A}$ macrophages. Li *et al*. put forth a possible explanation for this discrepancy: in IKK$\alpha^{-/-}$ knockouts, the absent IKKα proteins are no longer competing with IKKβ for NEMO binding locations, thereby allowing more IKKβ to homodimerize under NEMO [28]. These cells containing IKKβ-IKKβ -NEMO complexes have higher IκB kinase activity and thus less measured IκB than in IKK$\alpha^{A/A}$ mutants and normal cells that contain IKKα-IKKβ-NEMO complexes.

Our model as presented in this paper uses only two states to describe the state of a protein. This approximation suffices when the behavior of an inactivated protein is the same as that when that protein is absent. In this case, it is fine to associate both the absence and inactivation of the protein into state 0. However, in the case of IKKα, the effect of the protein's absence is different from that of its inactivation. Hence, full accuracy requires an additional state to encode for this level of detail. Accordingly, our model has included the pathway associated with IKKα's inactivation, and therefore, matches the observations of Lawrence *et al*'s inactivation model because their experimental method resulted in an inactivation of IKKα rather than its complete absence as in Li *et al*.

E. $IKK\beta^{-/-}$

Park *et al*. [29] used fetal liver-derived macrophages (FLDM) deficient in IKKβ to investigate the mechanism of macrophage survival under stimulus to TLR4 receptors. Some bacterial toxins, such as *Salmonella* AvrA, inhibit NF-κB while stimulating the TLR4 receptor, which was observed to result in the stimulation of macrophage apoptosis. This is mirrored in our model in Table II(f), where we see that p65 activity drops while the AP-1 proapoptotic family remains activated. This observation is along the same lines as that in Table II(b) where, in addition, we also tracked alterations in IκB activity under different exogenous stimuli conditions.

F. $NEMO^{-/-}$

Kim *et al*. [30] analyzed NEMO$^{-/-}$ Murine B cells. Because these mice die early in embryogenesis, they used an *in vitro* differentiation process to convert embryonic stem cells to B cells. They found that NEMO is not required for B cell development, but does affect its survival. Specifically, after an application of LPS for three days (+LPS) or mock stimulation for the control ($-$LPS), the wild-type B cells maintained population levels while the +LPS NEMO-deficient group declined in population. Oddly enough, however, the $-$LPS NEMO-deficient cell group also declined in similar proportions, which confounds the simple explanation of NF-κB stimulation from LPS increasing the cell apoptosis rate.

In our model, we see in Table II(g) that our NEMO$^{-/-}$ simulation actually shows an increase in p65 NF-κB activation levels with a constant AP-1 level, which would seem to indicate an increase in cell survival. This conflicts with the finding in Kim *et al*.; however, we see that the $\mathrm{IKK}\alpha = 1 \implies \mathrm{p65} = 0$ pathway in Table I is from [27], where the inhibition of p65 from IKKα is seen in macrophages. It is possible that this pathway is

different in developed B cells; indeed, if we remove this pathway from the model, then we see in Table II(h) that the p65 levels are unchanged in the NEMO-deficient model, meaning the model lacks sufficient information to predict any change in behavior for this knockout configuration. This reinforces a key assumption: the model is only as accurate as the pathway data used in its generation. For biological regulatory systems, it is typical to use pathways from different cell types and contexts; however, to achieve maximum model fidelity and prediction, it is necessary to obtain pathways from the same cell types that we wish to make predictions about.

G. $I\kappa B^{-/-}$

Klement *et al*. [31] and Beg *et al*. [32] investigated IκBα deficient murine embryonic fibroblasts, and under TNFα stimulation wildtype cells returned to low NF-κB levels rapidly whereas IκBα deficient cells remained high for up to two hours after withdrawing TNFα. Our model also exhibited this behavior with dynamic p65 levels after the removal of TNFα.

Also, [32] [in Fig. 7(c)] found that under LPS stimulation, IκBα deficient cells sustained higher amounts of NF-κB after 90 min. Again our model shows this in Table II(i).

Our model predicts an increase in constitutive (no exogenous stimulation) p65 expression under conditions and a decrease in p52 under both TNFα and LPS stimulation in Table II(i) and (j). Unfortunately, Klement and Beg found constitutive p65 levels to remain constant and p52 levels to be low in all experimental setups disagreeing with our model's predictions. Klement and Beg conclude that these findings can likely be explained by IkBβ, another member of the IkB family which was not included in our model, thus explaining the source of the discrepancy.

H. $RIP1^{-/-}$

Devin *et al*. [33] shows that RIP1 knockouts prevent TNFα stimulation from activating IKK and NF-kB. Our model also shows this in Table II(k). We assume the presence of cIAP in the cellular context as this RIP1 behavior was shown to be dependent on the presence of cIAP by recent work from Wong *et al*. [34].

IV. Conclusion

This paper presents a method to produce a regulatory network model using only minimal assumptions of predictor proteins and utilizing literature backed pathway information. The resulting networks are validated using biological knockout experiments from the literature. The use of minimal modeling assumptions together with the use of literature backed information results in a model built on a solid foundation of biological experimentation and suitable for further validation and refinement through comparison with high-throughput data and new pathway data as they become available. The linear nature of (marginal) biological pathways does not completely capture the nonlinear (multivariate) complex network level behaviors of reality; however, via the methods presented here, these pathways constrain network behavior in a way that could some day guide physician therapy and drug design aimed at complex pathways.

To extend this study, we will develop new techniques to leverage the data derived from high-throughput experiments to refine these pathway derived models. This will allow for networks that merge two sources of biological knowledge, prior knowledge and data, into models with better predictive power.

When going from pathways to networks, one must address two basic issues: inconsistencies in the pathway knowledge and incompleteness of that knowledge. The method of [8] addresses inconsistencies by imposing timing adjustments on the pathways. This does not contradict the pathway information because the pathways do not contain timing information; however, it requires a procedure external to the pathway knowledge and, although the imposition is reasonable, it is still outside the given knowledge. In our proposed method, inconsistencies are handled by introducing stochasticity into the model. While this requires a likelihood assumption, it does not impose timing information and it is done in a uniform manner so as to be minimally expansive.

Regarding incompleteness, which arises from unspecified entries in the Karnaugh maps, [8] keeps the original deterministic Boolean character of the pathways by creating an uncertainty class of deterministic Boolean networks, whereas here we interpret that regulatory uncertainty as randomness within a single network model. The different approaches lead to different types of models, both of which address the incompleteness of the pathway knowledge. Not only does the difference represent different modeling perspectives, it also engenders different approaches to working with the resulting models. For instance, if one wishes to intervene in network dynamics for the purpose of beneficially altering dynamic behavior, say, to reduce the likelihood of entering phenotypically undesirable states, then the single-network method proposed in the current paper allows a direct application of control theory in the framework of Markov decision processes to produce a control strategy that is optimal for the network relative to some given objective [35]. On the other hand, if one incorporates the pathway knowledge via an uncertainty class of networks, then the goal is to construct a control strategy this is optimally robust across the all networks in the uncertainty class relative to the objective [36], [37].

References

[1] E. Dougherty and A. Datta, "Genomic signal processing: Diagnosis and therapy," *IEEE Signal Process. Mag.*, vol. 22, no. 1, pp. 107–112, Jan. 2005.

[2] A. Datta and E. Dougherty, *Introduction to Genomic Signal Processing With Control*. Boca Raton, FL: CRC Press, 2007.

[3] M. Hecker, S. Lambeck, S. Toepfer, E. Van Someren, and R. Guthke, "Gene regulatory network inference: Data integration in dynamic models—A review," *Biosystems*, vol. 96, no. 1, pp. 86–103, 2009.

[4] P. Meyer, K. Kontos, F. Lafitte, and G. Bontempi, "Information-theoretic inference of large transcriptional regulatory networks," *EURASIP J. Bioinformatics Syst. Biol.*, vol. 2007, 2007.

[5] H. De Jong, "Modeling and simulation of genetic regulatory systems: A literature review," *J. Comput. Biol.*, vol. 9, no. 1, pp. 67–103, 2002.

[6] R. Albert, "Network inference, analysis, and modeling in systems biology," *The Plant Cell Online*, vol. 19, no. 11, p. 3327, 2007.

[7] B. Hayete, T. Gardner, and J. Collins, "Size matters: Network inference tackles the genome scale," *Molecular Syst. Biol.*, vol. 3, no. 1, 2007.

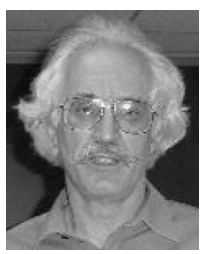
[8] R. Layek, A. Datta, and E. Dougherty, "From biological pathways to regulatory networks," *Mol. BioSyst.*, vol. 7, no. 3, pp. 843–851, 2011.
[9] G. Batt, D. Ropers, H. De Jong, J. Geiselmann, R. Mateescu, M. Page, and D. Schneider, "Validation of qualitative models of genetic regulatory networks by model checking: Analysis of the nutritional stress response in *Escherichia coli*," *Bioinformatics*, vol. 21 (Suppl. 1), p. i19, 2005.
[10] S. Kauffman, *The Origins of Order: Self Organization and Selection in Evolution*. New York: Oxford Univ. Press, 1993.
[11] M. Davidich and S. Bornholdt, "Boolean network model predicts cell cycle sequence of fission yeast," *PLoS One*, vol. 3, no. 2, p. 1672, 2008.
[12] B. Kuo, "*Digital Control Systems(Book)*," Champaign, IL: SRL Publishing Co., p. 572, 1977.
[13] S. Huang, "Gene expression profiling, genetic networks, and cellular states: An integrating concept for tumorigenesis and drug discovery," *J. Molecular Med.*, vol. 77, no. 6, pp. 469–480, 1999.
[14] M. Karnaugh, "The map method for synthesis of combinational logic circuits electrical engineering," Trans. AIEE, vol. 72, pp. 593–599, 1953.
[15] M. S. Hayden and S. Ghosh, "Signaling to NF-κB," *Genes Develop.*, vol. 18, no. 18, pp. 2195–2224, 2004.
[16] S. Ghosh, M. J. May, and E. B. Kopp, "Nf-κb and rel proteins: Evolutionarily conserved mediators of immune responses," *Annu. Rev. Immunology*, vol. 16, no. 1, pp. 225–260, 1998.
[17] S. Simmons, C. Fan, and R. Ramabhadran, "Cellular stress response pathway system as a sentinel ensemble in toxicological screening," *Toxicological Sci.: An official Jour. Soc. Toxicology*, vol. 111, no. 2, p. 202, 2009.
[18] M. Schmidt-Supprian, W. Bloch, G. Courtois, K. Addicks, A. Israel, K. Rajewsky, and M. Pasparakis, "NEMO/IKK [gamma]-deficient mice model incontinentia pigmenti," *Molecular cell*, vol. 5, no. 6, pp. 981–992, 2000.
[19] R. Cheong, A. Hoffmann, and A. Levchenko, "Understanding NF-κB signaling via mathematical modeling," *Molecular Syst. Biol.*, vol. 4, no. 1, 2008.
[20] S. Zhang, A. Kovalenko, G. Cantarella, and D. Wallach, "Recruitment of the IKK Signalosome to the p55 TNF Receptor: RIP and A20 Bind to NEMO (IKK [gamma]) upon Receptor Stimulation," *Immunity*, vol. 12, no. 3, pp. 301–311, 2000.
[21] M. Prendes, Y. Zheng, and A. Beg, "Regulation of developing B cell survival by RelA-containing NF-κ B complexes," *J. Immunol.*, vol. 171, no. 8, p. 3963, 2003.
[22] E. Dougherty, "Validation of gene regulatory networks: Scientific and inferential," *Brief. Bioinform.*, vol. 12, no. 3, p. 245, 2011.
[23] E. Dougherty and M. Bittner, *Epistemology of the Cell: A Systems Perspective on Biological Knowledge*, (IEEE Press Series on Biomedical Engineering series), 2011.
[24] S. Werner, D. Barken, and A. Hoffmann, "Stimulus specificity of gene expression programs determined by temporal control of IKK activity," *Science*, vol. 309, no. 5742, p. 1857, 2005.
[25] Q. Li, D. Antwerp, F. Mercurio, K. Lee, and I. Verma, "Severe liver degeneration in mice lacking the IB kinase 2 gene," *Science*, vol. 284, no. 5412, p. 321, 1999.
[26] Q. Li, Q. Lu, V. Bottero, G. Estepa, L. Morrison, F. Mercurio, and I. Verma, "Enhanced NF-κB activation and cellular function in macrophages lacking IκB kinase 1 (IKK1)," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 35, p. 12425, 2005.
[27] T. Lawrence, M. Bebien, G. Liu, V. Nizet, and M. Karin, "IKKα limits macrophage NF-κB activation and contributes to the resolution of inflammation," *Nature*, vol. 434, no. 7037, pp. 1138–1143, 2005.
[28] M. Hayden and S. Ghosh, "Shared principles in NF-[kappa] B signaling," *Cell*, vol. 132, no. 3, pp. 344–362, 2008.
[29] J. Park, F. Greten, A. Wong, R. Westrick, J. Arthur, K. Otsu, A. Hoffmann, M. Montminy, and M. Karin, "Signaling pathways and genes that inhibit pathogen-induced macrophage apoptosis–CREB and NF-[kappa] B as key regulators," *Immunity*, vol. 23, no. 3, pp. 319–329, 2005.
[30] S. Kim, R. La Motte-Mohs, D. Rudolph, J. Zúñiga-Pflücker, and T. Mak, "The role of nuclear factor-κB essential modulator (NEMO) in B cell development and survival," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 3, p. 1203, 2003.
[31] J. Klement, N. Rice, B. Car, S. Abbondanzo, G. Powers, P. Bhatt, C. Chen, C. Rosen, and C. Stewart, "Ikappabalpha deficiency results in a sustained NF-kappab response and severe widespread dermatitis in mice," *Molecular Cellular Biol.*, vol. 16, no. 5, p. 2341, 1996.
[32] A. Beg, W. Sha, R. Bronson, and D. Baltimore, "Constitutive NF-kappa b activation, enhanced granulopoiesis, and neonatal lethality in i kappa b alpha-deficient mice.," *Genes develop.*, vol. 9, no. 22, p. 2736, 1995.
[33] A. Devin, A. Cook, Y. Lin, Y. Rodriguez, M. Kelliher, and Z. Liu, "The distinct roles of TRAF2 and RIP in IKK activation by TNF-R1: TRAF2 recruits IKK to TNF-R1 while RIP mediates IKK activation," *Immunity*, vol. 12, no. 4, pp. 419–429, 2000.
[34] W. Wong, I. Gentle, U. Nachbur, H. Anderton, D. Vaux, and J. Silke, "RIPK1 is not essential for tNFR1-induced activation of NF-κb," *Cell Death Differ.*, vol. 17, no. 3, pp. 482–487, 2009.
[35] R. Pal, A. Datta, and E. Dougherty, "Optimal infinite-horizon control for probabilistic boolean networks," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2375–2387, Jun. 2006.
[36] R. Pal, A. Datta, and E. Dougherty, "Robust intervention in probabilistic boolean networks," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1280–1294, Mar. 2008.
[37] R. Pal, A. Datta, and E. Dougherty, "Bayesian robustness in the control of gene regulatory networks," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3667–3678, Sep. 2009.

Jason M. Knight (S'08) received the B.S. degree in biomedical engineering from the Texas A&M University, College Station, in 2009, where since then he has been working toward the Ph.D. degree in electrical engineering.

His current research interests include integrating heterogeneous sources of prior knowledge and integrating them into robust, optimal schemes for modeling, and inference of genetic regulatory networks.

Aniruddha Datta (F'09) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1991.

He is currently a Professor and holder of the J. W. Runyon, Jr. 35 Professorship II in the Department of Electrical and Computer Engineering, Texas A&M University, College Station. His research interest include adaptive control, robust control, PID control, and Genomic Signal Processing. He has authored or coauthored 5 books and more than 100 journal and conference papers on these topics.



Edward R. Dougherty (F'11) received the Ph.D. degree in mathematics from Rutgers University, New Brunswick, NJ.

He is a Professor in the Department of Electrical and Computer Engineering at Texas A&M University, College Station, where he holds the Robert M. Kennedy 26 Chair in Electrical Engineering and is the Director of the Genomic Signal Processing Laboratory. He is the Co-Director of the Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ. He is also with the Department of Bioinformatics and Computational Biology, University of Texas M. D. Anderson Cancer Center, Houston.

Dr. Dougherty is a fellow of SPIE, and has received the SPIE Presidents Award. He has been awarded the Doctor Honoris Causa by the Tampere University of Technology, Tampere, Finland.