Published in final edited form as:

IEEE Signal Process Mag. 2012 ; 29(1): 158–163. doi:10.1109/MSP.2011.943128.

Methods for Optimal Intervention in Gene Regulatory Networks

Nidhal Bouaynaya,

Systems Engineering Department, University of Arkansas at Little Rock

Roman Shterenberg, and

Department of Mathematics, University of Alabama at Birmingham

Dan Schonfeld

Department of Electrical and Computer Engineering, University of Illinois at Chicago

Nidhal Bouaynaya: nxbouaynaya@ualr.edu; Roman Shterenberg: shterenb@math.uab.edu; Dan Schonfeld: dans@uic.edu

The cell maintains its function via an elaborate network of interconnecting positive and negative feedback loops of genes and proteins that send different signals to a large number of pathways and molecules. These structures are referred to as *gene regulatory networks*, and their dynamics are used to understand the mechanisms and characteristics of biological cells as well as to search for possible remedies for various diseases such as cancer. Current research in cancer biology indicates that global, systemic molecular interactions are pivotal in understanding cellular dynamics and designing intervention strategies to combat genetic disease. In particular, most genetic diseases, such as cancer, are not caused by a single gene but rather by the interaction of multiple genes. Global, holistic approaches to the study of biological systems reveal the dynamic nature of cellular networks, which provide an important framework for drug discovery and design. The massive amounts of information that “omics” (e.g., genomics, proteomics, metabolomics) high-throughput sequencing technology generate mark a great leap forward in computational methods for analyzing and interpreting biological data. Currently, the major challenges are shifting toward optimal intervention strategies designed to affect the time evolution of gene activity in a desirable manner. One of the main aims of modern biological research is focused on intervening in biological cell dynamics to alter the gene regulatory network and avoid undesirable cellular states, e.g., metastasis.

In this column, we review recent results in gene regulatory network modeling and discuss various control mechanisms used to modify their cellular dynamics. We subsequently describe a new intervention strategy based on optimal perturbations, which force the network to converge to a desired steady-state distribution of gene regulation. The biological argument in support of the proposed framework is that steady-state distributions of gene regulatory networks determine the phenotype or the state of the cell development (for example, cell proliferation and apoptosis) [1]. Specifically, we investigate the existence, optimality, and robustness of perturbations that alter the dynamics of the network, leading to a desirable steady-state distribution. We subsequently present simulation results on the human melanoma gene regulatory network. Finally, we present a discussion of future trends and directions in control of gene regulatory networks. The ultimate goal is to develop engineering methods designed to intervene in the development of living organisms and transition cells from malignant states into benign forms.

GENE REGULATORY NETWORK MODELS

Network models of gene interactions serve the dual purpose of identifying organizational and dynamic principles as well as providing an exploratory framework for the development

of computational tools to study biological systems. In particular, understanding the dynamic behavior of gene regulatory networks is essential to advance our knowledge of disease, develop modern therapeutic methods, and identify targets in the cell needed to reach a desired goal. Therefore, major work has focused on building models of gene regulatory networks by inferring functional relationships among genes from gene expression profiles.

Computational models and methods for genetic network inference should account for uncertainties inherent to biological systems: stochastic fluctuations of molecular processes, incomplete knowledge, and the noisy nature of measurements. Subsequently, probabilistic models have been the most successful in elucidating the nature of gene interactions within the cell [2]. In particular, Markov chain models have been shown to accurately mimic the dynamic behavior of gene networks [2]. The Markov chain model encompasses several network class models, including the most widely adopted probabilistic Boolean networks (PBNs) and dynamic Bayesian networks (DBNs) [3]. The PBN is a stochastic extension of the standard Boolean network that incorporates probabilistic rule-based dependencies between its nodes, the genes. Bayesian networks (BNs) are directed acyclic graphs that represent dependencies between variables in a probabilistic model. DBNs relax the acyclicity constraint by taking into account the evolution of expression data over time. DBNs are able to capture several other often-used modeling frameworks, such as hidden Markov models (and its variants) and Kalman filter models, as its special cases. The two models are related, as Lähdesmäki et al. [3] showed, PBNs and a certain subclass of DBNs can represent the same joint probability distribution over their common variables. Therefore, an advantage of the proposed framework for intervention within Markovian gene regulatory networks is that it can be applied to a large class of network models, including PBNs and DBNs.

INTERVENTION STRATEGIES

The ultimate objective of gene regulatory network modeling and analysis is to use the network to design effective intervention strategies for affecting its dynamics in such a way as to avoid undesirable cellular states or phenotypes. As futuristic gene therapeutic interventions, various control strategies have been proposed to alter gene interactions in a desirable way. Even though the developed interventions remain, so far, as sheer theoretical investigations, such alterations may be biologically possible by the introduction of a drug or exposure to certain radiations that alter the extant behavior of the cell. In this context, the synergy between theoretical investigation and experimental validation is essential to establish an effective plan that will ultimately lead to the development of novel treatment and clinical decision making in genetic research.

Current interventions within gene regulatory networks can be grouped into three main approaches: 1) introduce external controls, which specify the interventions on control genes by optimizing a specific cost function [4]; 2) develop heuristic control policies based on certain dynamic properties of the network [5], [6]; and 3) alter the state transition structure of the network and, consequently, its long-run behavior. This last type of intervention is also referred to as *structural intervention* [1], [7], [5].

The first strategy determines a policy, i.e., a rule at each decision time point, which controls certain genes in the network by minimizing a given cost function. Two frameworks are considered: finite and infinite horizon controls [4]. Both control policies can be found as the solutions to the optimal stochastic control problems associated with their corresponding Bellman optimality equations. The external control requires knowledge of the target genes to be used as control variables as well as the cost function to be minimized. Moreover, the optimal policy is obtained through an iterative procedure that is computationally expensive

even for small-size networks. In addition to these drawbacks, the finite-horizon control may not change the long-run behavior of the network as it is applied over a finite-time window.

To alleviate the computational burden of the optimal external control, reduction techniques have been proposed that delete either genes or states [6]. However, deletion of network components reduces its size at the expense of information loss. Alternative solutions were found in various heuristic interventions, which also use external variables to specify interventions on control genes [5], [6]. In [6], a greedy stationary control policy using mean first passage times (MFPTs) of the Markov chain was proposed. The MFPT control policy is based on the intuition that the time to reach desirable states or leave undesirable states should be increased. Although the MFPT is closely related to the steady-state distribution, the MFPT control policy does not directly rely on the shift of the steady-state distribution. Heuristic control policies, which use the shift of stationary mass as criterion, have been proposed in [5].

Whereas the optimal external control and its approximations (heuristics) consist of policies that (recursively) alter control genes to optimize certain objective functions, structural intervention proposes to alter the dynamics governing the network to shift its steady-state mass to favorable cellular states. The motivation is that these states may represent different phenotypes, or cellular functional states, such as quiescence, and one would want to decrease the probability that the network will end up in an undesirable set of states [1]. Shmulevich et al. [1] used genetic algorithms to alter the steady-state probabilities of certain states. An analytical study of steady-state distributions for structurally perturbed PBNs was presented in [7]. The analysis, however, focuses on rank-one perturbations, and the extension of the method to higher-rank perturbations is iterative and computationally very expensive.

A general solution to the problem of shifting the steady-state mass of gene regulatory networks, modeled as Markov chains, has been recently advanced in [8]. The proposed framework, which can be viewed as a generalization of the work in [7], formulates optimal intervention in gene regulation as a solution to an inverse perturbation problem and demonstrates that the solution, if it exists, is unique, globally optimum, and can be computed efficiently using standard convex optimization methods. The inverse perturbation problem addresses the following problem: “Given a Markov chain characterized by its probability transition matrix P_0 and given a desired steady-state distribution π_d , find a (optimal) perturbation matrix C that forces the chain $P_0 + C$ to converge to the desired distribution π_d as its unique steady-state distribution.” Observe that, in contrast to perturbation theory, which finds the new steady-state distribution given a known perturbation, the inverse perturbation framework aims at finding a perturbation that forces the network to transition to a known desired steady-state mass. The analytical solution to the optimal inverse problem provides a minimally perturbed Markov chain characterized by a steady-state distribution corresponding to the desired probability mass. The perturbation can be thought of as reshaping the attractor landscape to have a unique desired stationary distribution, with the entire state-space as its basin of attraction. An illustration is presented in Figure 1.

OPTIMAL PERTURBATION

We consider a gene regulatory network with m genes, where the expression level of each gene is quantized to l values. The dynamic behavior of this network can be represented as a finite-state Markov chain described by a probability transition matrix P_0 of size $n = l^m$. A row probability vector $\mu^t = (\mu_1, \dots, \mu_n)$ is called a *stationary distribution*, or a *steady-state distribution*, for P_0 if $\mu^t P_0 = \mu^t$. Because P_0 is stochastic (i.e., its rows sum up to one), the existence of stationary distributions is guaranteed.

Let π_0 denote the undesirable steady-state distribution of P_0 . We wish to alter this distribution by linearly perturbing the probability transition matrix P_0 . Specifically, we consider the perturbed stochastic matrix $P = P_0 + C$, where C is a zero row-sum perturbation matrix (the zero row-sum condition is necessary to ensure that the perturbed matrix P is a proper probability matrix). Let us denote by π_d the desired stationary distribution. We seek to find a zero row-sum perturbation matrix C such that the perturbed matrix P admits π_d as a steady-state distribution.

FEASIBILITY PROBLEM

The set of perturbation matrices C , which force the network to converge towards π_d as its unique stationary distribution, satisfy the following constraints:

$$\begin{aligned} & \text{(i) } \pi_d^t = \pi_d^t (P_0 + C), \\ & \text{(ii) } C\mathbf{1} = \mathbf{0}, \\ & \text{(iii) } P_0 + C \geq 0, \\ & \text{(iv) } \text{SLEM}(P_0 + C) < 1, \end{aligned} \quad (1)$$

where constraint (iii) denotes an element-wise inequality. SLEM stands for second largest eigenvalue modulus, where the eigenvalues are counted taking into account their multiplicity. In particular, (iv) implies that eigenvalue 1 is simple. Along with the positivity of the desired distribution, constraint (iv) is equivalent to ergodicity. Constraints (ii) and (iii) ensure that the perturbed matrix P is a proper probability transition matrix. Let $\mathcal{F}$ denote the feasible set of perturbation matrices, i.e.,

$$\begin{aligned} \mathcal{F} = \{ & C \in \mathbb{R}^{n \times n} : \pi_d^t \\ & = \pi_d^t (P_0 + C), C\mathbf{1} = \mathbf{0}, \\ & P_0 + C \geq 0, \text{SLEM}(P_0 + C) < 1 \}. \end{aligned} \quad (2)$$

The feasible set $\mathcal{F} \neq \emptyset$ because $(\mathbf{1}\pi_d^t - P_0) \in \mathcal{F}$. In fact, the second largest eigenvalue modulus (SLEM) $(\mathbf{1}\pi_d^t) = 0 < 1$, and it is easy to check that $(\mathbf{1}\pi_d^t - P_0)$ satisfies conditions (i)–(iii). Therefore, the feasible set is not empty, and we can find at least one feasible perturbation matrix, which forces the network to converge toward a desired steady-state distribution.

OPTIMAL PERTURBATION

To answer the question, “Which perturbation(s) in $\mathcal{F}$ is optimal?” we need to adopt optimality criteria, where an objective function is optimized subject to the constraints in (1). We propose optimality criteria, which may clinically translate into minimizing potential adverse effects caused by the intervention strategy. Specifically, we focus on minimization of the change in the structure of the network and maximization of the convergence rate toward the steady-state distribution. We will therefore investigate the following criteria for optimal perturbation control:

- Reduce the overall level of change before and after intervention as measured by the energy of the network dynamics, i.e., minimize the energy of change between the original and perturbed transition matrices as characterized by the Frobenius-norm of the perturbation matrix.
- Increase the rate of convergence of the network to the desired steady-state distribution, thus minimizing the time needed to reach the desired steady-state distribution.

MINIMAL-ENERGY PERTURBATION—The minimal-energy perturbation matrix is defined as the feasible perturbation with minimum Frobenius-norm. Analytically, the minimal-energy perturbation is obtained as the solution of the following optimization problem:

$$\text{Minimize } \|C\|_F \text{ subject to } C \in \mathcal{F}, \quad (3)$$

where $\|\cdot\|_F$ denotes the Frobenius-norm given by $\|C\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n c_{ij}^2$. Because the Frobenius-norm is strictly convex, the optimization problem in (3) has, at most, one minimizer. In general, the optimal solution belongs to the closure of the set, $\overline{\mathcal{F}} = \mathcal{D} = \{C \in \mathbb{R}^{n \times n} : \pi_d^t = \pi_d^t(P_0 + C), C\mathbf{1} = \mathbf{0}, P_0 + C \geq \mathbf{0}\}$. That is, $\mathcal{D}$ is the set of perturbation matrices satisfying conditions (i)–(iii) only. Denote by C_* the minimum Frobenius-norm perturbation matrix over the closure $\mathcal{D}$. We know that C_* exists and is unique because the set $\mathcal{D}$ is convex and closed. Moreover, C_* can be computed efficiently as the solution of a semidefinite programming algorithm [8]. It can be shown that C_* is the optimum solution of the optimization problem in (3) if $C_* \in \mathcal{F}$ [9]. Otherwise, we can find a feasible perturbation that is arbitrarily close to C_* [9], i.e., given $\delta > 0$, we can find $C \in \mathcal{F}$ such that $\|C - C_*\|_F < \delta$.

The optimization problem formulated in (3) using the Frobenius-norm may also be casted using a different matrix norm. In this case, the problem interpretation will also be different. For instance, the evaluation of the Frobenius-norm leads to an energy interpretation of the network, whereas using the L_1 -norm, for instance, would lead to a sparse perturbation matrix. The sparsity criterion would translate into the minimal number of changes introduced to the dynamics of the initial network to force it to attain the desired steady-state distribution. However, because the L_1 -norm is not strictly convex, the optimal L_1 -norm solution is not necessarily unique.

FASTEST CONVERGENCE RATE PERTURBATION—Another clinically viable optimality criterion is to select the perturbation that yields the fastest convergence rate to the desired steady-state distribution. The Markov chain, which models the network dynamics, converges to its unique steady-state distribution if and only if its SLEM is strictly smaller than 1. In this case, the convergence rate is given by the SLEM. The smaller the SLEM, the faster the Markov chain converges to its stationary distribution. The optimal fastest convergence rate perturbation, C_R^* is therefore obtained as the feasible perturbation with the smallest SLEM, i.e., C_R^* is the solution to the following optimization problem

$$\begin{aligned} &\text{Minimize SLEM}(P_0 + C) \\ &\text{subject to } C \in \mathcal{F}. \end{aligned} \quad (4)$$

For a general (nonsymmetric) matrix, about the only characterization of the eigenvalues is the fact that they are the roots of the characteristic polynomial. In particular, the objective function in (4) is not necessarily convex, and thus the optimization problem is not convex. However, an evident solution to the fastest convergence rate perturbation problem is given by $C_R^* = \mathbf{1}\pi_d^t - P_0$. The optimal $\text{SLEM}(P_0 + C_R^*) = 0$. That is, the perturbation C_R^* seems to force the network to reach the desired steady-state in a single jump. Observe that the fastest convergence rate perturbed network matrix is equal to the desired limiting matrix $\mathbf{1}\pi_d^t$.

MINIMAL-ENERGY AND FASTEST CONVERGENCE-RATE PERTURBATIONS TRADEOFF—The fastest convergence rate perturbation may result in a large energy

deviation between the original and perturbed networks. Similarly, the minimal-energy perturbation may lead to a slow convergence (hence slow recovery) to the desired steady state. We will, therefore, investigate the tradeoffs between minimal-energy and fastest convergence rate criteria. For this purpose, we consider the family of matrices, parameterized by s , along the line between P_E^* , the minimal-energy perturbed matrix, and $\mathbf{1}\pi_d^t$, the fastest convergence rate perturbed matrix, $P(s) = (1-s)P_E^* + s\mathbf{1}\pi_d^t$, where $0 \leq s \leq 1$. The parameterized line $P(s)$ can be thought of as a continuous transformation of P_E^* into $\mathbf{1}\pi_d^t$. One can easily check that the parameterized perturbations $C(s) = P(s) - P_0$ are feasible perturbations, i.e., $C(s) \in \mathcal{C}$ for all $0 \leq s \leq 1$. When $s = 0$, we have the minimal-energy perturbed matrix, and when $s = 1$, we obtain the fastest convergence rate perturbed matrix. When $0 < s < 1$, there is a tradeoff between the energy value and the rate of convergence. Specifically, it is shown in [8] that when s increases, the SLEM of the perturbed matrix decreases, and hence the convergence (toward the desired state) is faster. On the other hand, the energy deviation between the original and perturbed networks increases with s . In other words, the faster the network converges toward the desired steady state, the higher the energy between the initial and perturbed networks.

ROBUSTNESS OF OPTIMAL PERTURBATION

The (optimal) inverse perturbation framework requires knowledge of the probability transition matrix of the network. The probability transition matrix can be estimated from the gene expression profiles [2]. However, errors made during the measurement, data extraction, and model parameter selection will propagate to the inference of the probability transition matrix and thus affect the actual success of the designed control. An efficient intervention strategy must be robust to data and estimation errors. The minimal-energy inverse perturbation is shown to be robust to errors in the original probability transition matrix in the sense that the estimation error of the optimal perturbation is bounded by the estimation error of the probability transition matrix [10]. Analytically, let the estimated probability transition matrix $\hat{P}_0$ be given by $\hat{P}_0 = P_0 + \delta P_0$, where δP_0 is a zero-row sum matrix representing noisy and missed data and estimation errors in P_0 . Then the estimated optimal perturbation matrix, $\hat{C}^*$, satisfies $\hat{C}^* = C^* + \delta C^*$, where C^* is the optimal perturbation matrix and δC^* is a zero-row sum error matrix satisfying $\|\delta C^*\|_F \leq \|\delta P_0\|_F$.

GENE REGULATION IN MELANOMA CELLS

We apply the inverse perturbation control to the melanoma gene regulatory network, which is one of the most studied gene regulatory networks in the literature [2]. The abundance of mRNA for the gene WNT5A was found to be highly discriminating between cells with properties typically associated with high versus low metastatic competence. Furthermore, it was found that an intervention that blocked the Wnt-5a protein from activating its receptor, the use of an antibody that binds the Wnt-5a protein could substantially reduce Wnt-5A's ability to induce a metastatic phenotype [4]. This suggests a control strategy that reduces WNT5A's action in affecting biological regulation. A seven-gene probabilistic Boolean network model of the melanoma network containing the genes WNT5A, pirin, S100P, RET1, MART1, HADHB, and STC2 was derived in [2]. Note that the human melanoma Boolean network consists of $2^7 = 128$ states ranging from $00 \dots 0$ to $11 \dots 1$, where the states are ordered as WNT5A, pirin, S100P, RET1, MART1, HADHB, and STC2, with WNT5A and STC2 denoted by the most significant bit (MSB) and least significant bit (LSB), respectively.

We consider the (fictitious) desired steady-state distribution where the probability of the states having WNT5A upregulated is 10^{-4} and the probability of the other states, which correspond to WNT5A downregulated is chosen randomly such that the total probability

mass is equal to one [see Figure 2(c)]. Observe that the states 0–63 have WNT5A downregulated and hence are desirable states, as compared to states 64–127, which have WNT5A upregulated and hence are undesirable. The probability transition matrices of the human melanoma networks corresponding to the original and perturbed networks are portrayed in Figure 2(a) and (b), respectively. The matrix plots are obtained using the function `MatrixPlot` in `MATHEMATICA`. The color of entries varies from white to red, corresponding to the values of the entries in the range of zero to one. Note that the controlled and desired steady-state distributions are identical. Moreover, we have $\text{SLEM}(P) < 1$. Therefore, the minimal Frobenius-norm perturbation is the optimal perturbation, which forces the network to converge toward the desired stationary distribution, independently of the initial state of the network.

FUTURE TRENDS AND DIRECTIONS

The solution to the inverse perturbation problem characterizes the optimal state probability transition matrix that yields the optimal perturbation of the regulatory network and transitions to the desired steady-state. To reach the full impact of the optimal intervention on gene regulation in biological systems, we must introduce changes in the cell that induce the optimal transition matrix. The state probability transition matrix is derived from gene expression profiles [2]. The probability transition matrix is related to the actual gene network by observing that the probability law describing the genes' dynamics can be obtained as the marginal distribution of the state transition probabilities

$$\Pr(g_i = x_i | g_1 \dots, g_m) = \sum_{\tilde{x}_i} \Pr(g_1 = x_1, \dots, g_m = x_m | g_1 \dots, g_m), \quad (5)$$

where $\tilde{x}_i$ denotes the set of all x_j 's except x_i , i.e., $\tilde{x}_i = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m\}$. In particular, one can show that small perturbations of the probability transition matrix lead to small perturbations of the genes' dynamics, thus validating the minimal-perturbation criterion.

Future work will investigate changes in the cell that induce the optimal perturbed transition matrix. In particular, we will focus on determining the optimal biological design rules needed to induce the optimal intervention strategy while limiting ourselves to biologically viable design rules that rely on modern methods for molecular control in cellular systems. The biological rules should identify whether a specific gene will excite (upregulate) or inhibit (downregulate) a target gene. Implementation of such rules and associated probabilities can be achieved using modern biological methods for molecular control in cellular systems.

References

1. Shmulevich I, Dougherty ER, Zhang W. Control of stationary behavior in probabilistic Boolean networks by means of structural intervention. *J Biol Syst.* 2002; 10(4):431–445.
2. Kim S, Li H, Dougherty ER, Cao N, Chen Y, Bittner M, Suh EB. Can Markov chain models mimic biological regulation? *J Biol Syst.* 2002; 10(10):337–357.
3. Lähdesmäki H, Hautaniemi S, Shmulevich I, Yli-Harjao O. Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks. *Signal Processing.* 2006; 86(4):814–834. [PubMed: 17415411]
4. Datta A, Pal R, Choudhary A, Dougherty E. Control approaches for probabilistic gene regulatory networks—What approaches have been developed for addressing the issue of intervention? *IEEE Signal Processing Mag.* Jan; 2007 24(1):54–63.

5. Qian X, Ivanov I, Ghaffari N, Dougherty ER. Intervention in gene regulatory networks via greedy control policies based on long-run behavior. *BMC Syst Biol.* Jun; 2009 3(1):61, 61–77. [PubMed: 19527511]
6. Vahedi G, Faryabi B, Chamberland J, Datta A, Dougherty E. Intervention in gene regulatory networks via a stationary mean-first-passage-time control policy. *IEEE Trans Biomed Eng.* Oct; 2008 55(10):2319–2331. [PubMed: 18838357]
7. Qian X, Dougherty ER. Effect of function perturbation on the steady-state distribution of genetic regulatory networks: Optimal structural intervention. *IEEE Trans Signal Processing.* Oct; 2008 52(10):4966–4976.
8. Bouaynaya N, Shterenberg R, Schonfeld D. Inverse perturbation for optimal intervention in gene regulatory networks. *Bioinformatics.* 2011; 27(1):103–110. [PubMed: 21062762]
9. Bouaynaya, N.; Rasheed, M.; Shterenberg, R.; Schonfeld, D. Intervention in general topology gene regulatory networks. *Proc. IEEE Int. Workshop Genomic Signal Processing and Statistics*; Dec 2011;
10. Bouaynaya, N.; Shterenberg, R.; Schonfeld, D. Robustness of inverse perturbation for discrete event control. *Proc. IEEE Int. Conf. Engineering in Medicine and Biology Society*; Aug 2011;

FIG 1.

Inverse perturbation problem of dynamic networks characterizes the optimal reshaping of the attractor landscape of the network such that the basin of attraction consists of a unique desired distribution for the entire state space.

FIG 2.

Optimal perturbation of the human melanoma gene regulatory network: (a) the probability transition matrix of the original melanoma network P_0 , (b) the minimal-energy perturbed probability transition matrix P_E^* , (c) the fastest convergence rate perturbed matrix, and (d) the steady-state distributions corresponding to the original (red line), desired (blue line), and minimal-perturbation energy controlled (green line) human melanoma gene regulatory network.