Optimal Perturbation Control of General Topology Molecular Networks

Nidhal Bouaynaya, *Member, IEEE*, Roman Shterenberg, and Dan Schonfeld, *Fellow, IEEE*

***Abstract*—In this paper, we develop a comprehensive framework for optimal perturbation control of dynamic networks. The aim of the perturbation is to drive the network away from an undesirable steady-state distribution and to force it to converge towards a desired steady-state distribution. The proposed framework does not make any assumptions about the topology of the initial network, and is thus applicable to general-topology networks. We define the optimal perturbation control as the minimum-energy perturbation measured in terms of the Frobenius-norm between the initial and perturbed probability transition matrices of the dynamic network. We subsequently demonstrate that there exists at most one optimal perturbation that forces the network into the desirable steady-state distribution. In the event where the optimal perturbation does not exist, we construct a family of suboptimal perturbations, and show that the suboptimal perturbation can be used to approximate the optimal limiting distribution arbitrarily closely. Moreover, we investigate the robustness of the optimal perturbation control to errors in the probability transition matrix, and demonstrate that the proposed optimal perturbation control is robust to data and inference errors in the probability transition matrix of the initial network. Finally, we apply the proposed optimal perturbation control method to the Human melanoma gene regulatory network in order to force the network from an initial steady-state distribution associated with melanoma and into a desirable steady-state distribution corresponding to a benign cell.**

***Index Terms*—Control, dynamical systems, gene regulatory networks, Markov chains, perturbation.**

I. Introduction

ELUCIDATION of the interactions between molecular structures in biological organisms can provide valuable insights into human diseases, including cancer [1]. Spurred by advances in molecular profiling technology, computational models of cellular networks have been sought, and a variety of algorithms to infer the structure of molecular networks have been proposed and evaluated [2]–[7]. From a translational perspective, molecular networks provide the ultimate framework to develop of genomic-based therapy and treatment. The main challenge is to design optimal intervention strategies that rely on molecular network models in order to avoid undesirable cellular states, e.g., metastasis. Modification of the molecular network may be biologically viable by perturbation of the expression level of target genes. Such perturbations have been shown to occur as a result of the introduction of a chemical substance (e.g., acid or RNAi) or exposure to radiation (e.g., light) that may alter the extant behavior of the cell [8]. Recent methods developed in the field of biotechnology have demonstrated the ability to perturb, silence or activate the expression level of genes in a desirable manner [8].

Despite all of the work that has been devoted to the inference of genomic and proteomic regulatory networks from profile datasets, little effort has been devoted thus far towards the control of molecular networks. The difficulty stems from the fact that traditional control theory, developed for engineering systems, is not readily applicable to biological systems. Traditional control schemes rely on an exogenous control signal provided to input nodes, and used to minimize the total cost of the system [9]. In biology, however, cost functions of molecular systems are unknown and there are no obvious input variables that can be used to control the system. Even if target genes were to be used as control variables, they may not be able to control all of the genes in the network and thus cannot guarantee controllability of the regulatory network [10]. Nonetheless, by assuming knowledge of a cost function and of target genes, Datta *et al.* [11] and Faryabi *et al.* [12] derived an optimal control policy, i.e., a rule at each time instant used to control the target genes in order to minimize the cost function. Different choices of the cost function lead to distinct control strategies, in terms of efficiency, computational complexity and robustness [13], [14]. Furthermore, the effect of the external control is limited to the duration of application of the control. That is, the network will remain controlled only in the presence of the control inputs. Once the control inputs have been removed, the network will rely on its original dynamics and will no longer be controlled, and the system is thus likely to transition to an undesirable state. In biology, however, we are interested in the steady-state behavior of gene regulatory networks, which has been associated with phenotypes such as cell proliferation and apoptosis [15].

An efficient control strategy for biological networks should, therefore, alter the dynamics governing the molecular network in order to shift its steady-state mass to a favorable cellular state [16]. This limitation suggests a different strategy for biological intervention from traditional control. Shmulevich *et al.* [16] relied on genetic algorithms to modify the steady-state distribution of undesirable states. In [17], the impact of function per-

Manuscript received December 28, 2011; revised October 30, 2012 and December 14, 2012; accepted December 28, 2012. Date of publication January 18, 2013; date of current version March 12, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yufei Huang. This project is supported by Award Number R01GM096191 from the National Institute Of General Medical Sciences (NIH/NIGMS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of General Medical Sciences or the National Institutes of Health.

N. Bouaynaya is with the Department of Systems Engineering, University of Arkansas, Little Rock, AR 72204 USA (e-mail: nxbouaynaya@ualr.edu).

R. Shterenberg is with the Department of Mathematics, University of Alabama, Birmingham, AL 35294 USA (e-mail: shterenb@math.uab.edu).

D. Schonfeld is with the Department of Electrical and Computer Engineering, University of Illinois, Chicago, IL 60607 USA (e-mail: dans@uic.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2241054

turbations in Boolean networks was explored. However, the approach proposed was limited to singleton attractors. Moreover, the algorithms developed were computationally complex due to the requirement to determine the state changes before and after perturbations. An analysis of steady-state distributions for structurally perturbed probabilistic Boolean networks (PBNs) was presented in [18]. This analysis, however, focused on rank-one perturbations, and a proposed extension to higher-rank perturbations relied on an iterative and computationally expensive method.

A general method for network perturbation was introduced in [19]. The proposed approach was based on the design of a perturbation that changes the steady-state distribution of the network from its original undesirable steady-state distribution to a desirable one. However, the method is limited to ergodic, i.e., irreducible and aperiodic networks. Furthermore, although the method presented in [19] ensures that the desired state is a possible steady-state for the perturbed network, it does not guarantee that the perturbed network will converge towards the desired steady-state distribution. In fact, a network may possess a desired steady-state distribution, but fail to converge to the desired state. In systems biology, however, we are not only interested in changing the dynamical landscape of the molecular network, but must also force the network to converge to the desired steady-state distribution.

In this paper, we present a comprehensive framework to address optimal intervention in general-topology, not necessarily ergodic, networks and the convergence of the networks towards a desired steady-state distribution. The proposed framework of optimal perturbation control in general topology networks can be applied to any discrete-time system, which can be modeled by a finite-state Markov chain model. Examples include queuing networks, resource allocation, social and biological networks, and machine replacement. The remainder of the paper is organized as follows: In Section II, we investigate the feasibility of perturbation control of general-topology networks. In Section III, we formulate and derive the optimal perturbation algorithm. Robustness of the proposed approach to optimal perturbation control is investigated in Section IV. In Section V, we present simulation results for both a synthetic network as well as a biological network (namely, the Human melanoma gene regulatory network). Finally, in Section VI, we provide a brief summary and discussion of the main results of the paper.

In this paper, we consider real variables. We use $\mathbb{R}$ to denote the set of real numbers. Scalars are denoted by lower case letters, e.g., s, t. Vectors in $\mathbb{R}^n$ are denoted by bold letters, numbers, or lower-case Greek letters, e.g., $\mathbf{1}$, $\mathbf{x}$, $\boldsymbol{\pi}$, where $\mathbf{1}$ denotes a vector all of whose components are equal to one. $\mathbf{x}^t$ denotes the transpose of the vector $\mathbf{x}$. Matrices in $\mathbb{R}^{m\times n}$ are denoted by bold capital letters or upper-case Greek letters, e.g., $\boldsymbol{C}$, $\boldsymbol{P}$, $\boldsymbol{\Lambda}$. $\boldsymbol{I}$ stands for the identity matrix.

II. Feasibility Problem

We consider a gene regulatory network with m genes $g_1, \ldots, g_m$, where the expression level of each gene is quantized to l values. The expression levels of all genes in the network defines the state vector of the network. The dynamics of this network can be represented by a finite-state homogeneous Markov chain described by a probability transition matrix $\boldsymbol{P}_0$ of size $n = l^m$ [20]. The probability transition matrix encapsulates the one-step conditional probabilities of the genes thus indicating the likelihood that the network will evolve from one state vector to another. We do not assume any particular structure on the initial network topology $\boldsymbol{P}_0$.

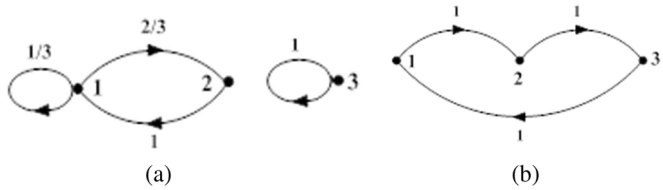

Fig. 1. (a) The graph of a reducible Markov chain. (b) The graph of an irreducible but periodic Markov chain.

Definition 1: A row probability vector $\boldsymbol{\mu}^t = (\mu_1, \ldots, \mu_n)$ is called a stationary distribution, or a steady-state distribution, for $\boldsymbol{P}$ if $\boldsymbol{\mu}^t\boldsymbol{P} = \boldsymbol{\mu}^t$.

Because $\boldsymbol{P}$ is stochastic (i.e., its rows sum up to unity), 1 is an eigenvalue of $\boldsymbol{P}$, and, therefore, $\boldsymbol{P}$ has at least one stationary distribution. The chain is *irreducible* if its state space is a single communicating class; in other words, if every state is reachable from every other state. If $\boldsymbol{P}$ is irreducible, it has a unique stationary distribution $\boldsymbol{\pi}$ and $\boldsymbol{\pi}$ is strictly positive [21]. If $\boldsymbol{P}$ is irreducible and aperiodic, it is called *ergodic*. For an ergodic probability transition matrix $\boldsymbol{P}$, we have convergence towards the unique, strictly positive, steady-state distribution, in the following sense,

$$\lim_{n\to\infty} \boldsymbol{P}^n = \mathbf{1}\boldsymbol{\pi}^t. \tag{1}$$

Equation (1) states that for any initial state distribution $\boldsymbol{\mu}_0$, we have $\lim_{n\to\infty} \boldsymbol{\mu}_0^t\boldsymbol{P}^n = \boldsymbol{\pi}^t$. That is, the network converges to the stationary distribution $\boldsymbol{\pi}$ from any initial state distribution or the basin of attraction of $\boldsymbol{\pi}$ is the entire state-space. A Markov chain can fail to converge for two reasons (or combinations thereof).

1) The chain is reducible, as in the example shown in Fig. 1(a), where

$$\boldsymbol{P} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In this case, the set of all stationary distributions is the convex hull spanned by $(0,0,1)^T$ and $\left(\frac{1}{2},\frac{1}{3},\frac{1}{6}\right)^T$. In particular, the matrix $\boldsymbol{P}$ does not converge towards a rank-one matrix.

2) The chain is irreducible but periodic, as in the example shown in Fig. 2(b), where

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

In this case, even though the chain admits a unique stationary distribution, $\boldsymbol{\pi} = \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)^T$, $\boldsymbol{P}$ does not converge to $\mathbf{1}\boldsymbol{\pi}^t$.

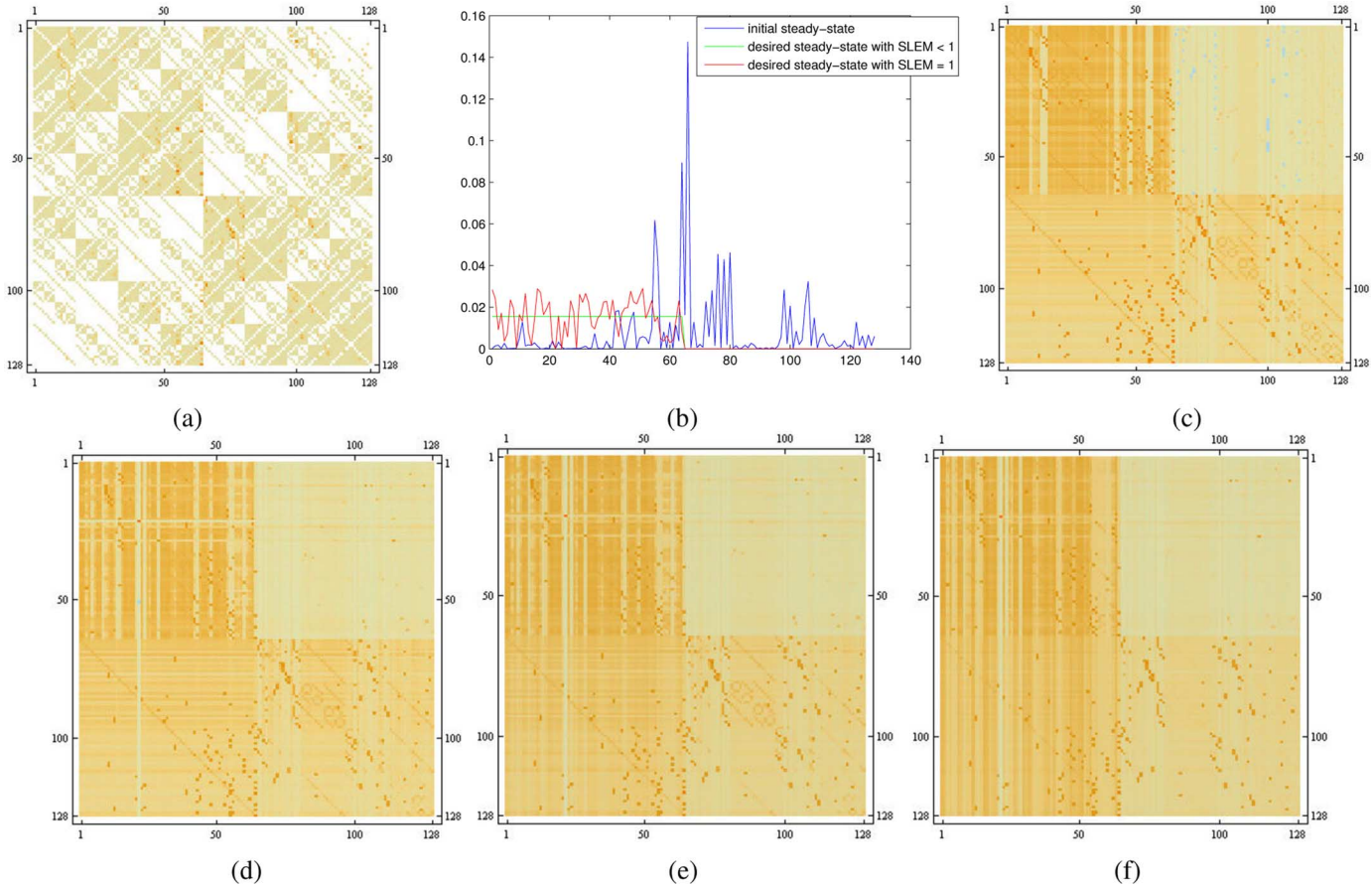

Fig. 2. Optimal inverse perturbation of the Human melanoma gene regulatory network. (a) The probability transition matrix, $\boldsymbol{P}_0$, of the melanoma gene regulatory network. (b) The initial steady-state distribution (blue) and two different desired steady-state distributions, $\boldsymbol{\pi}_d^1$ (green) and $\boldsymbol{\pi}_d^2$ (red), which correspond to a downregulation of the gene WNT5A. (c) The optimal perturbed matrix $\boldsymbol{P}_1^*$ corresponding to the steady-state distribution $\boldsymbol{\pi}_d^1$ (green). $\boldsymbol{P}_1^*$ converges to $\boldsymbol{\pi}_d^1$. (d) The optimal perturbed matrix $\boldsymbol{P}_2^*$ corresponding to the steady-state distribution $\boldsymbol{\pi}_d^2$ (red). $\boldsymbol{P}_2^*$ does not converge to $\boldsymbol{\pi}_d^2$. (e) The perturbed matrix $\boldsymbol{P}_\epsilon = (1-\epsilon)\boldsymbol{P}_2^* + \epsilon \mathbf{1}\boldsymbol{\pi}_d^t$ converging towards $\boldsymbol{\pi}_d^2$. (f) The perturbed matrix $\boldsymbol{P}_a = \boldsymbol{P}_2^* + (1-a)(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}_2^*)$ converging towards $\boldsymbol{\pi}_d^2$ at the same rate as $\boldsymbol{P}_1^*$ converges towards $\boldsymbol{\pi}_d^1$.

Let us consider an initial molecular network, modeled as a Markov chain process, with probability transition matrix $\boldsymbol{P}_0$. Unlike the work in [19], which assumed an ergodic probability transition matrix, we do not impose any particular structure on $\boldsymbol{P}_0$ or the underlying Markov process. Actually, we intuitively suspect large molecular networks to be reducible and/or periodic, hence non-ergodic. Therefore, the initial network has at least one steady-state distribution. Suppose one initial steady-state distribution is undesirable, e.g., reflects a disease cellular state. Our goal is to design an optimal perturbation matrix $\boldsymbol{C}$ such that the perturbed matrix $\boldsymbol{P}$ converges to the desired steady-state distribution $\boldsymbol{\pi}_d$ as its unique steady-state distribution. In other words, we propose to alter the dynamical landscape of the network by replacing all initial stationary distributions by the unique desirable steady-state. We perturb the matrix $\boldsymbol{P}_0$ as

$$\boldsymbol{P} = \boldsymbol{P}_0 + \boldsymbol{C}, \tag{2}$$

where $\boldsymbol{C}$ is a zero-row sum perturbation matrix. The zero row-sum condition is necessary to ensure that the perturbed matrix $\boldsymbol{P}$ is stochastic. Let us denote by $\boldsymbol{\pi}_d$ the desired stationary distribution. The following Lemma provides a necessary and sufficient condition for a stochastic matrix to converge towards its steady-state distribution.

Lemma 1: Consider a stochastic matrix $\boldsymbol{P}$. Let $\boldsymbol{\pi}$ denote a stationary distribution of $\boldsymbol{P}$. Then, we have

$$\lim_{n\to\infty} \boldsymbol{P}^n = \mathbf{1}\boldsymbol{\pi}^t \iff \mathrm{SLEM}(\boldsymbol{P}) < 1, \tag{3}$$

where SLEM denotes the second largest eigenvalue magnitude, where the eigenvalues are counted with their algebraic multiplicity. In other words, the stochastic matrix converges towards its steady-state distribution if and only if 1 is a simple eigenvalue of $\boldsymbol{P}$ and all other eigenvalues have magnitude strictly less than 1.

It follows from Lemma 1 that a perturbation matrix, which forces the network to converge towards the desired steady-state distribution, must satisfy the following four conditions:

(i) $\boldsymbol{\pi}_d^t(\boldsymbol{P}_0 + \boldsymbol{C}) = \boldsymbol{\pi}_d^t$
(ii) $\boldsymbol{C}\mathbf{1} = \mathbf{0}$
(iii) $\boldsymbol{P}_0 + \boldsymbol{C} \geq 0$
(iv) $\mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}) < 1$

Condition (i) states that $\boldsymbol{\pi}_d$ is a stationary distribution of $\boldsymbol{P}_0 + \boldsymbol{C}$ (not necessarily unique). Condition (iv) establishes that the stationary distribution is unique and the perturbed matrix converges towards it. Conditions (ii) and (iii) ensure that the perturbed matrix $\boldsymbol{P}$ is a proper probability transition matrix, i.e., it is stochastic and elementwise non-negative. Let $\mathcal{F}$ denote the

feasible set of perturbation matrices, i.e., $\mathcal{F}$ is the set of matrices $\boldsymbol{C}$ satisfying conditions (i) through (iv),

$$\mathcal{F} = \{\boldsymbol{C} \in \mathbb{R}^{n\times n} : \boldsymbol{\pi}_d^t(\boldsymbol{P}_0 + \boldsymbol{C}) = \boldsymbol{\pi}_d^t, \boldsymbol{C}\mathbf{1} = \mathbf{0}, \boldsymbol{P}_0 + \boldsymbol{C} \geq 0, \mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}) < 1\}. \quad (4)$$

Observe that $\boldsymbol{C}_0 = \mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}_0 \in \mathcal{F}$. It is straightforward to check that $\boldsymbol{C}_0$ satisfies conditions (i), (ii), and (iii). Moreover, $\mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}_0) = \mathrm{SLEM}(\mathbf{1}\boldsymbol{\pi}_d^t) = 0 < 1$. In particular, the feasible set $\mathcal{F} \neq \emptyset$, and therefore, there exists at least one perturbation, $\boldsymbol{C}$, which forces the network to converge towards the desired steady-state. The following proposition provides a characterization of the feasible set of perturbations.

Proposition 1: Given a stochastic matrix $\mathbf{P}_0$ and given a desired probability vector $\boldsymbol{\pi}_d$ not proportional to $\mathbf{1}$, consider the basis $\mathcal{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ formed by the vectors $\mathbf{x}_1 = \mathbf{1}$, $\mathbf{x}_2 = (\mathbf{1}, \boldsymbol{\pi}_d) \perp \boldsymbol{\pi}_d$, i.e., $\mathbf{x}_2$ is a linear combination of $\mathbf{1}$ and $\boldsymbol{\pi}_d$ that is orthogonal to $\boldsymbol{\pi}_d$, and $\mathbf{x}_i \perp (\mathbf{1}, \boldsymbol{\pi}_d)$ for $i \geq 3$. If $\boldsymbol{\pi}_d$ is proportional to $\mathbf{1}$, then define $\boldsymbol{x}_1 = \mathbf{1}$ and $\boldsymbol{x}_j \perp \boldsymbol{x}_1$ for $j \geq 2$. Let $\mathbf{A}$ be the representation matrix of basis $\mathcal{X}$ in the canonical basis. Then, the perturbed matrix $\mathbf{P} = \mathbf{P}_0 + \mathbf{C}$ is similar to a block form matrix, $\mathbf{P}_{\mathcal{X}}$, i.e.,

$$\boldsymbol{P} = \mathbf{A}\mathbf{P}_{\mathcal{X}}\mathbf{A}^{-1}, \quad (5)$$

where

$$\mathbf{P}_{\mathcal{X}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & * & * & * \\ \vdots & * & * & * \\ 0 & * & * & * \end{pmatrix}, \quad (6)$$

and the $*$'s are any real value such that $\boldsymbol{P} \geq 0$ elementwise, and the magnitude of the maximum eigenvalue of the $(n-1)\times(n-1)$ submatrix of $*$'s is strictly less than unity.

For instance, a feasible solution is obtained when all values $*$ are equal to zero. In this case, $\boldsymbol{P} = \mathbf{1}\boldsymbol{\pi}_d^t$. From Proposition 1, we can see that there are infinitely many perturbations, which force the network to converge towards the desired steady-state distribution $\boldsymbol{\pi}_d$. All such perturbations are plausible intervention strategies, and can be used to drive the network towards the desired steady-state at equilibrium. At this point, a question arises "Which feasible perturbation(s) are optimal?" To answer this question, we need to define an optimality criterion. In this paper, we consider the minimum energy constraint in order to minimize the overall "energy" of change between the initial and perturbed networks.

III. Optimal Perturbation Control

We define the "energy" of a network as the Frobenius norm of its probability transition matrix. The Frobenius norm of $\boldsymbol{C} = \{c_{ij}\}$ is defined as $\|\boldsymbol{C}\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n c_{ij}^2 = \mathrm{Tr}\,(\boldsymbol{C}^T\boldsymbol{C})$, where $\mathrm{Tr}(\boldsymbol{X})$ denotes the trace of matrix $\boldsymbol{X}$. The minimum-energy perturbation, which drives the network towards the desired steady-state, is the solution to the following optimization problem

$$\text{Minimize } \|\boldsymbol{C}\|_F^2 \text{ subject to } \boldsymbol{C} \in \mathcal{F}, \quad (7)$$

where $\mathcal{F}$ is the feasible set defined in (4). Since the Frobenius norm is strictly convex, there exists at most one solution to the problem in (7). In general, the optimal solution belongs to the closure, $\bar{\mathcal{F}} \supseteq \mathcal{F}$, of the feasible set $\mathcal{F}$, where $\bar{\mathcal{F}}$ is given by

$$\bar{\mathcal{F}} = \{\boldsymbol{C} : \boldsymbol{\pi}_d^t(\boldsymbol{P}_0 + \boldsymbol{C}) = \boldsymbol{\pi}_d^t, \boldsymbol{C}\mathbf{1} = \mathbf{0}, \boldsymbol{P}_0 + \boldsymbol{C} \geq 0\}. \quad (8)$$

Denote by $\boldsymbol{C}^*$ the minimum Frobenius norm perturbation matrix over the closure $\bar{\mathcal{F}}$. We know that $\boldsymbol{C}^*$ exists and is unique because the set $\bar{\mathcal{F}}$ is convex and closed. Moreover, $\boldsymbol{C}^*$ can be computed efficiently as the solution of a semi-definite programming algorithm [19].

A geometric characterization of the optimal perturbation can be obtained as follows:

Proposition 2: Let $\boldsymbol{C}_0 = \mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}_0$, and consider the set Γ defined as

$$\Gamma = \{\boldsymbol{V} \in \mathbb{R}^{n\times n} : \boldsymbol{V}\mathbf{1} = \mathbf{0}, \boldsymbol{V}^t\boldsymbol{\pi}_d = \mathbf{0}, \boldsymbol{V} \geq -\mathbf{1}\boldsymbol{\pi}_d^t\}. \quad (9)$$

Then, the optimal perturbation matrix (over $\bar{\mathcal{F}}$) is given by

$$\boldsymbol{C}^* = \boldsymbol{C}_0 - \boldsymbol{C}_0^{\Gamma}, \quad (10)$$

where $\boldsymbol{C}_0^{\Gamma}$ denotes the unique projection of $\boldsymbol{C}_0$ onto the convex set Γ[1].

This geometric characterization will be useful in the robustness analysis of the optimal perturbation (see Section IV).

In what follows, we show that $\boldsymbol{C}^*$ is also the unique optimal solution of the problem in (7) if $\boldsymbol{C}^* \in \mathcal{F}$.

Proposition 3: Let $\boldsymbol{C}^* = \arg\min_{\boldsymbol{C}\in\bar{\mathcal{F}}} \|\boldsymbol{C}\|_F^2$. Then, $\boldsymbol{C}^* \in \bar{\mathcal{F}}$. Moreover, if $\boldsymbol{C}^* \in \mathcal{F}$, then it is the unique optimal solution of (7).

Two important points can be drawn from Proposition 3. First, the existence of the optimal perturbation, which drives the network towards the desired steady-state distribution, is independent of the initial structure of $\boldsymbol{P}_0$ (i.e., whether the initial network is ergodic or not). In particular, this result generalizes the work in [19]. Furthermore, even if $\boldsymbol{P}_0$ was ergodic, the optimal perturbation may result in a non-ergodic perturbed network. That is, the optimal perturbation may alter the structure and connection properties of the initial network. Second, the existence of the optimal perturbation depends on the specific values of $\boldsymbol{P}_0$ and $\boldsymbol{\pi}_d$. For instance, if $\boldsymbol{\pi}_d$ has zero entries, then the perturbed probability transition matrix $\boldsymbol{P} = \boldsymbol{P}_0 + \boldsymbol{C}^*$ will be reducible, i.e., not fully communicating, and may not converge towards $\boldsymbol{\pi}_d$.

In the case that the optimal perturbation matrix does not exist (due, for instance, to the particular structure of $\boldsymbol{P}_0$ and $\boldsymbol{\pi}_d$), we can still get arbitrarily close to the optimal solution by considering a sequence $\boldsymbol{C}_n \in \mathcal{F}$ which converges towards $\boldsymbol{C}^*$. The following proposition provides a construction of such a sequence.

Proposition 4: Assume that $\boldsymbol{C}^* \notin \mathcal{F}$, i.e., $SLEM(\boldsymbol{P}_0 + \boldsymbol{C}^*) = 1$. Consider the family of matrices described by

$$\boldsymbol{C}_n = (1 - \epsilon_n)\boldsymbol{C}^* + \epsilon_n(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}_0), \quad (11)$$

where $0 < \epsilon_n \leq 1$ is a sequence converging to zero, i.e., $\lim_{n\to\infty} \epsilon_n = 0$. Then, we have

1) $\boldsymbol{C}_n \in \mathcal{F}, \forall\, n \in \mathrm{IN}$.

[1]Note that projection, here, refers to minimal distance to the set. Because Γ is not a vector subspace, we cannot define (orthogonal) projection onto Γ. However, we can determine the closest point in Γ to a given point in $\mathbb{R}^{n\times n}$.

2) $\lim_{n\to\infty} \boldsymbol{C}_n = \boldsymbol{C}^*$
3) $\|\boldsymbol{C}_n\|_F > \|\boldsymbol{C}^*\|_F, \forall\, n \in \mathbb{N}$.

In particular, the perturbation matrix, $\boldsymbol{C} = (1-\epsilon)\boldsymbol{C}^* + \epsilon(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}_0)$, where $\epsilon > 0$ is any small number, approximates the optimal perturbation, and forces the network out of its undesirable steady-state distribution and into the desirable one.

From Lemma 1, the convergence of the perturbed network towards the desired steady-state is guaranteed by the condition $\mathrm{SLEM}(\boldsymbol{P}) < 1$. On the other hand, the second largest eigenvalue modulus defines also the rate of convergence towards the steady-state distribution [21]. The smaller the SLEM, the faster the convergence towards the steady-state. In particular, if the SLEM is close to unity (while being strictly smaller to 1), the perturbed network will converge extremely slowly to the desired steady-state, and the designed control will become unpractical. Therefore, we propose to replace the constraint $\mathrm{SLEM}(\boldsymbol{P}) < 1$ by $\mathrm{SLEM}(\boldsymbol{P}) \leq a$, for some $a < 1$. This new constraint has two advantages: First, it avoids the slow convergence scenario by making sure that the convergence rate is at least equal to $a < 1$. Second, it ensures the existence of the optimal perturbation because the feasible set becomes closed. The price paid, however, is that the new optimal perturbation has a higher energy than the optimal solution of (7).

Let us consider the new feasible set

$$\mathcal{D} = \{\boldsymbol{C} : \boldsymbol{\pi}_d^t(\boldsymbol{P}_0 + \boldsymbol{C}) = \boldsymbol{\pi}_d^t, \boldsymbol{C}\mathbf{1} = \mathbf{0}, \boldsymbol{P}_0 + \boldsymbol{C} \geq \mathbf{0}, \mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}) \leq a\}, \quad (12)$$

where $a < 1$. Then, the optimization problem

$$\text{Minimize } \|\boldsymbol{C}\|_F^2 \text{ subject to } \boldsymbol{C} \in \mathcal{D} \quad (13)$$

admits a global solution because the Frobenius norm is strictly convex and the feasible set $\mathcal{D}$ is closed. However, unlike $\bar{\mathcal{F}}$, $\mathcal{D}$ is not convex because the SLEM function is not convex for nonsymmetric matrices. Therefore, the optimal solution may not be unique. Moreover, the optimization problem in (13) cannot be solved using convex minimization methods or solvers [22]. Nevertheless, we can find an explicit sub-optimal solution if we restrict the search space to the line between the minimum energy perturbed matrix, $\boldsymbol{P}^* = \boldsymbol{P}_0 + \boldsymbol{C}^*$ and the limiting matrix $\mathbf{1}\boldsymbol{\pi}_d^t$, i.e., we consider the following parameterized family of matrices

$$\boldsymbol{P}(s) = (1-s)\boldsymbol{P}^* + s\mathbf{1}\boldsymbol{\pi}_d^t, \quad 0 \leq s \leq 1. \quad (14)$$

Equation (14) can be thought of as a continuous transformation of P^* into $\mathbf{1}\boldsymbol{\pi}_d^t$. The perturbation matrix $\boldsymbol{C}(s) = \boldsymbol{P}(s) - \boldsymbol{P}_0$ is then given by

$$\boldsymbol{C}(s) = \boldsymbol{C}^* + s(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}^*), \quad 0 \leq s \leq 1. \quad (15)$$

It is easy to verify that the family $\boldsymbol{C}(s) \in \bar{\mathcal{F}}$. The minimum energy solution over the parameterized set of perturbation matrices $\boldsymbol{C}(s)$ can be computed using the following Lemma from [19].

Lemma 2 [19]: Consider the family of parameterized matrices in (14). We have

1) $SLEM(\boldsymbol{P}(s)) = (1-s)SLEM(\boldsymbol{P}^*)$.
2) $\|\boldsymbol{C}(s)\|_F$ is an increasing function of s.

In particular, observe that for $0 < s \leq 1$, the perturbed family of matrices $\boldsymbol{P}(s)$ has SLEM < 1, and therefore, converges towards its unique steady-state distribution. Lemma 2 states that the SLEM of the perturbed matrix decreases when s increases, and hence the convergence (towards the desired steady-state) is faster. On the other hand, the norm of the perturbation matrix, and hence the energy deviation between the original and perturbed networks, increases as a function of s. Therefore, the minimal energy perturbation over the family $\{\boldsymbol{C}(s)\}_{0\leq s\leq 1}, \boldsymbol{C}_a$, is obtained when $\mathrm{SLEM}(\boldsymbol{P}(s)) = a$. Explicitly, we have

$$\boldsymbol{C}_a = \boldsymbol{C}^* + (1 - \frac{a}{\mathrm{SLEM}(\boldsymbol{P}^*)})(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}^*). \quad (16)$$

If the optimal perturbation does not exist, i.e., $\mathrm{SLEM}(P^*) = 1$, we have

$$\boldsymbol{C}_a = \boldsymbol{C}^* + (1-a)(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}^*). \quad (17)$$

IV. Robustness of the Optimal Perturbation Control

In practice, the probability transition matrix of the initial network, $\boldsymbol{P}_0$, is estimated using expression data [20]. Errors made during data extraction, feature selection, and network inference will propagate and impact the actual success of the designed control. An efficient intervention approach must, therefore, possess some degree of robustness or insensitivity to data and estimation errors.

We assume that the estimated probability transition matrix $\hat{\boldsymbol{P}}_0$ is given by

$$\hat{\boldsymbol{P}}_0 = \boldsymbol{P}_0 + \delta\boldsymbol{P}_0, \quad (18)$$

where $\delta\boldsymbol{P}_0$ is a zero-row sum matrix representing noisy and missed data and estimation errors in $\boldsymbol{P}_0$. We show that the norm of the error in the optimal perturbation matrix is bounded by the norm of the error in $\boldsymbol{P}_0$. We first present the following Lemma.

Lemma 3: Consider a vector subspace $\mathcal{E}$, equipped with an inner product norm $\|\cdot\|$, and a convex subset $\mathcal{C} \subseteq \mathcal{E}$. Let x_1, x_2 be two points in $\mathcal{E}$, and p_1, p_2 be their respective closest points in $\mathcal{C}$. Then, we have

$$\|p_1 - p_2\| \leq \|x_1 - x_2\|. \quad (19)$$

Proposition 5: The estimated optimal perturbation matrix, $\hat{\boldsymbol{C}}^*$, satisfies $\hat{\boldsymbol{C}}^* = \boldsymbol{C}^* + \delta\boldsymbol{C}^*$, where $\boldsymbol{C}^*$ is the optimal perturbation matrix and $\delta\boldsymbol{C}^*$ is an error zero-row sum matrix satisfying

$$\|\delta\boldsymbol{C}^*\|_F \leq \|\delta\boldsymbol{P}_0\|_F. \quad (20)$$

That is the norm of the error in the optimal inverse perturbation matrix is bounded by the norm of the error in $\boldsymbol{P}_0$, and the optimal perturbation control is robust to data and inference errors.

V. Simulation Results

The Markov probability transition matrix, describing the dynamics of the network at the state level, 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:

$$\Pr(g_i = x_i | g_1 \ldots, g_m) = \sum_{\tilde{x}_i} \Pr(g_1 = x_1, \ldots, g_m = x_m | g_1 \ldots, g_m), \quad (21)$$

where $\tilde{x}_i$ denotes the set of all x_j's except x_i; i.e., $\tilde{x}_i = \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m\}$. In order to capture the dynamics of the gene network, a "wiring rule" is considered in [20] such that the expression level of each gene at the next step is predicted by the expression levels of the genes at the current step. We define the gene network matrix, G, as the matrix whose entries are the conditional probabilities of the individual genes expression levels given the current network state, i.e., given the expression levels of all other genes. We order the columns of G such that the first l columns indicate the probabilities of gene $g_1 = 0, g_1 = 1, \ldots, g_1 = l - 1$, respectively, given the network states; the next l columns provide the probabilities of gene $g_2 = 0, 1, \ldots, l$ given the network states, and so on. For instance, for a binary quantization ($l = 2$), we have $G(1,1) = \Pr(g_1 = 0 | g_1 = 0, g_2 = 0) = \Pr(g_1 = 0|00)$ and $G(1,2) = \Pr(g_1 = 1|00)$. Formally, the gene network matrix, for a binary quantization, is defined as

$$G(i,j) = \Pr(g_{k_j} = x_{k_j} | \text{ network state } i),$$

where

$$k_j = \begin{cases} j/2, & \text{if } j \text{ is even;} \\ (j+1)/2, & \text{if } j \text{ is odd.} \end{cases}$$

$$x_{k_j} = \begin{cases} 0, & \text{if } j \text{ is even;} \\ 1, & \text{if } j \text{ is odd.} \end{cases}$$

We first illustrate how the proposed optimal perturbation control alters the dynamics of the gene network through the following example.

A. Synthetic Example

Consider a two-gene network, g_1, g_2, where the expression level of each gene is quantized to 0 (downregulated) and 1 (upregulated). Hence, the state-space of the network has 4 states: 00, 01, 10, and 11. Let us assume that the probability transition matrix of this network is given by

$$\boldsymbol{P}_0 = \begin{pmatrix} 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.3 & 0.2 \\ 0.3 & 0.3 & 0 & 0.4 \\ 0.1 & 0.3 & 0.6 & 0 \end{pmatrix}. \quad (22)$$

The gene network matrix can be computed using (22) as

$$\boldsymbol{G}_0 = \begin{pmatrix} 0 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.8 & 0.2 \\ 0.6 & 0.4 & 0.3 & 0.7 \\ 0.4 & 0.6 & 0.7 & 0.3 \end{pmatrix}. \quad (23)$$

Observe that gene g_1 is always upregulated when the network is at state 00. We have $\Pr(g_1 = 0|00) = \Pr(g_1 = 0, g_2 = 0|00) + \Pr(g_1 = 0, g_2 = 1|00) = 0$. Hence $\Pr(g_1 = 1|00) = 1$. Also, observe that there are no state feedback loops, i.e., the diagonal of the p.t.m. $\boldsymbol{P}_0$ has zero entries. This network converges towards the steady-state distribution

$$\boldsymbol{\pi}_0 = [0.2162, 0.1809, 0.3276, 0.2753]^t. \quad (24)$$

Assume that the state 10, which has steady-state mass equal to 0.3276 corresponds to an undesirable cellular state, and all other states are equally desirable. We, therefore, wish to design an optimal perturbation, which forces the network to converge towards the desired steady-state

$$\boldsymbol{\pi}_d = [1/3, 1/3, 0, 1/3]^t. \quad (25)$$

The optimally perturbed p.t.m. is computed as

$$\boldsymbol{P}^* = \begin{pmatrix} 0.1444 & 0.2444 & 0.0000 & 0.6111 \\ 0.5778 & 0.1778 & 0.0000 & 0.2444 \\ 0.3000 & 0.3000 & 0.0000 & 0.4000 \\ 0.2778 & 0.5778 & 0.0000 & 0.1444 \end{pmatrix}. \quad (26)$$

The gene network dynamics after control is, therefore, given by

$$\boldsymbol{G}^* = \begin{pmatrix} 0.3888 & 0.6111 & 0.1444 & 0.8555 \\ 0.7556 & 0.2444 & 0.5778 & 0.4222 \\ 0.6000 & 0.4000 & 0.3000 & 0.7000 \\ 0.8556 & 0.1444 & 0.2778 & 0.7222 \end{pmatrix}. \quad (27)$$

Observe that, in the controlled network, gene 1 may become downregulated if the network is at state 00: we have $\Pr(g_1 = 0|00) = \boldsymbol{G}^*(1,1) = 0.3888$. On the other hand, the control did not alter the genes' expressions given the network state 10, i.e., the third rows of $\boldsymbol{G}_0$ and $\boldsymbol{G}^*$ are identical. The control, however, introduced state feedback loops (the diagonal of the perturbed p.t.m. $\boldsymbol{P}^*$ is not identically zero). Finally, observe that even though the optimally perturbed p.t.m is reducible (state 10 is not reachable), it converges towards the desired steady-state distribution $\boldsymbol{\pi}_d$ as its unique steady-state distribution.

B. Human Melanoma Gene Regulatory Network

We consider the Human melanoma (skin cancer) gene regulatory network [23]. 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 Wnt5a protein from activating its receptor, the use of an antibody that binds the Wnt5a protein, could substantially reduce Wnt5A's ability to induce a metastatic phenotype [11], [18], [23]. 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 [24]. The Human melanoma Boolean network consists of $2^7 = 128$ states ranging from $00\cdots0$ to $11\cdots1$, 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.

Because the aim is to downregulate the WNT5A gene, the states from 64 to 127, which correspond to WNT5A upregulated, should have near zero steady-state mass. In our simulations, we consider two different desired steady-state distribu-

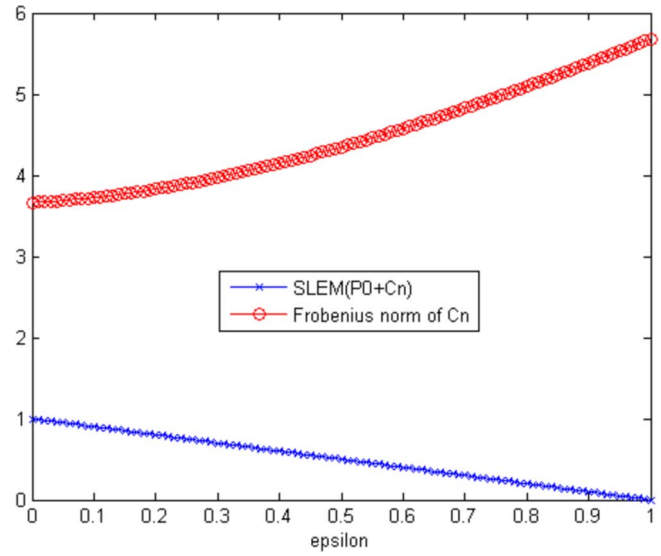

Fig. 3. $\mathrm{SLEM}(\boldsymbol{P}_0+\boldsymbol{C}_n)$ versus ϵ_n (blue), and $\|\boldsymbol{C}_n\|_F$ versus ϵ_n (red), where $\boldsymbol{C}_n$ is given by (11).

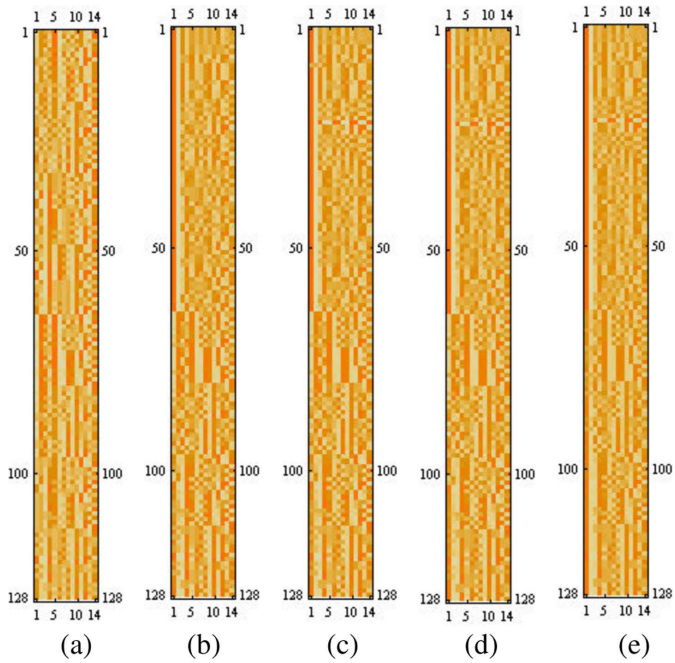

Fig. 4. Optimal inverse perturbation of the Human melanoma gene regulatory network. (a) The Human melanoma gene network matrix, $\boldsymbol{G}_0$. (b) The optimal melanoma gene network matrix, $\boldsymbol{G}_1^*$ corresponding to the steady-state distribution $\boldsymbol{\pi}_d^1$. (c) The optimal melanoma gene network matrix, $\boldsymbol{G}_2^*$ corresponding to the steady-state distribution $\boldsymbol{\pi}_d^2$. (d) The optimal melanoma gene network matrix, $\boldsymbol{G}_\epsilon$ corresponding to the p.t.m. $\boldsymbol{P}_\epsilon$ (in Fig. 2(e)) for $\epsilon = 0.1$. (e) The optimal melanoma gene network matrix, $\boldsymbol{G}_a$ corresponding to the p.t.m. P_a (in Fig. 2(f)).

tions $\boldsymbol{\pi}_d^1$ and $\boldsymbol{\pi}_d^2$, shown in Fig. 2(b). The first distribution, $\boldsymbol{\pi}_d^1$, assigns probability 10^{-4} to the states having WNT5A upregulated and a uniform mass equal to 0.015525 to the other states. The second distribution, $\boldsymbol{\pi}_d^2$, also assigns a uniform mass of 10^{-4} to the undesirable states but assigns random probabilities to the other states such that the total probability mass is equal to 1. The first and second steady-state distributions are plotted in blue and red, respectively, in Fig. 2(b). The corresponding optimal perturbed transition matrices, $\boldsymbol{P}_1^*$ and $\boldsymbol{P}_2^*$, are depicted in Fig. 2(c) and (d), respectively. The original transition matrix, $\boldsymbol{P}_0$, is shown in Fig. 2(a). The matrix plots are obtained using the function *MatrixPlot* in MATHEMATICA. They provide a visual representation of the values of elements in the matrix. The color of entries varies from white to red corresponding to the values of the entries in the range of 0 to 1. We have $\mathrm{SLEM}(\boldsymbol{P}_1^*) < 1$ and $\mathrm{SLEM}(\boldsymbol{P}_2^*) = 1$. Therefore, the optimal perturbed network with p.t.m. $(\boldsymbol{P}_1^*)$ converges towards the desired stationary distribution $\boldsymbol{\pi}_d^1$, whereas there exists no optimal perturbation, which forces the network to converge towards $\boldsymbol{\pi}_d^2$. However, from Proposition 4, we can design perturbation matrices that are arbitrarily close to the optimal solution.

We now consider the desired steady-state distribution $\boldsymbol{\pi}_d^2$, which corresponds to a $\mathrm{SLEM}(\boldsymbol{P}_0+\boldsymbol{C}^*) = 1$ and hence an optimally perturbed matrix $\boldsymbol{P}_2^*$ which does not converge towards the desired steady-state. Proposition 4 states that the corresponding sequence of perturbation matrices $\boldsymbol{C}_n$, given by (11), correspond to ergodic perturbed matrices, which converge towards $\boldsymbol{\pi}_d^2$. In Fig. 3, we plotted $\mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}_n)$ and $\|\boldsymbol{C}_n\|_F$ versus ϵ_n. Observe that the SLEM is a decreasing function, whereas the Frobenius norm increases with ϵ_n. In particular, given a $\delta > 0$, there exists $\epsilon_n > 0$ such that $\|\boldsymbol{C}_n - \boldsymbol{C}_*\| < \delta$, and $\boldsymbol{C}_n \in \mathcal{F}$. Therefore, $\boldsymbol{C}_n$ can be considered as suboptimal solutions to the inverse perturbation problem in (7). Fig. 2(e) shows the p.t.m. $\boldsymbol{P}_\epsilon = (1-\epsilon)\boldsymbol{P}_2^* + \epsilon \mathbf{1}\boldsymbol{\pi}_d^t$ for $\epsilon = 0.1$, and Fig. 2(f) shows $\boldsymbol{P}_a = \boldsymbol{P}_2^* + (1-a)(\mathbf{1}\boldsymbol{\pi}_d^t - \boldsymbol{P}_2^*)$ for $a = SLEM(\boldsymbol{P}_1^*)$. From (17), $\boldsymbol{P}_a$ converges towards $\boldsymbol{\pi}_d^2$ at the same rate as $\boldsymbol{P}_1^*$ converges towards $\boldsymbol{\pi}_d^1$.

TABLE I
FROBENIUS DISTANCES BETWEEN THE PROBABILITY TRANSITION MATRICES IN FIG. 2

Frobenius distance	$\boldsymbol{P}_0$	$\boldsymbol{P}_1^*$	$\boldsymbol{P}_2^*$	$\boldsymbol{P}_\epsilon$	$\boldsymbol{P}_a$
$\boldsymbol{P}_0$	0	3.4329	3.6634	3.7269	4.1695
$\boldsymbol{P}_1^*$		0	1.3778	1.3638	1.9413
$\boldsymbol{P}_2^*$			0	0.3965	1.6364
$\boldsymbol{P}_\epsilon$				0	1.2398
$\boldsymbol{P}_a$					0

TABLE II
FROBENIUS DISTANCES BETWEEN THE GENE NETWORK MATRICES IN FIG. 4

Frobenius distance	$\boldsymbol{G}_0$	$\boldsymbol{G}_1^*$	$\boldsymbol{G}_2^*$	$\boldsymbol{G}_\epsilon$	$\boldsymbol{G}_a$
$\boldsymbol{G}_0$	0	9.6005	9.6001	9.8520	11.1158
$\boldsymbol{G}_1^*$		0	2.4268	2.5442	4.4130
$\boldsymbol{G}_2^*$			0	0.9307	3.8406
$\boldsymbol{G}_\epsilon$				0	2.9099
$\boldsymbol{G}_a$					0

The corresponding gene network matrices are plotted in Fig. 4. Tables I and II show the Frobenius distances between the probability transition matrices and the gene network matrices, respectively. In particular, observe that a "small" perturbation of the p.t.m. leads to a "small" perturbation in the corresponding gene network. This can be seen as follows: If the probability transition matrix $\boldsymbol{P}_0$ is perturbed linearly with a zero-row sum matrix $\boldsymbol{C} = \{\epsilon_{i,j}\}_{1\le i,j\le n}$, then conditional probability of each gene $\Pr(g_i = x_i | g_1, \ldots, g_m)$ is perturbed linearly by $\sum_{j\in J} \epsilon_{hj}$, where h is the index of the state vector $[g_1, \ldots, g_m]$ and J is an interval isomorphic to $\{1, 2, \ldots, \frac{n}{l}\}$.

In particular, "small" perturbations $\epsilon_{ij} \ll 1$ of the probability transition matrix that satisfy the zero-row sum condition $\sum_{j=1}^{n} \epsilon_{hj} = 0$, lead to "small" perturbations of the genes' dynamics.

The MATLAB and MATHEMATICA codes are posted online at *http://syen.ualr.edu/nxbouaynaya/TSP%202012.html*. Also, the numerical values of all output matrices in Figs. 2 and 4 can be downloaded in Excel format from the same website.

VI. Conclusion

In this paper, we presented a comprehensive framework for optimal perturbation control of general-topology networks. The aim of perturbation control is to perturb the network in such a way that it will drive the network away from an undesirable steady-state distribution and into a desirable one. We proved that there are infinitely many perturbations, which can serve as control strategies and achieve the aim of perturbation control. We defined the optimal perturbation as the minimum Frobenius-norm perturbation that minimizes the energy between the probability transition matrices of the initial and perturbed networks. We demonstrated that there exists at most one solution to the optimal perturbation control problem. The existence of an optimal perturbation control depends both on the initial network dynamics as well as the desired steady-state distribution. In the event that an optimal perturbation control does not exist, we constructed a family of suboptimal perturbations, which approximate the optimal limiting distribution arbitrarily closely. Moreover, we investigated the robustness of optimal perturbation control to errors in the initial probability transition matrix, and showed that the proposed perturbation control method is robust to data and inference errors in the probability transition matrix of the initial network. The proposed optimal perturbation control in general-topology networks can be applied to any system modeled as a discrete-time homogeneous Markov chain in order to reach a desired steady-state distribution. Examples include computer networks and social networks. We applied the proposed optimal perturbation control to the Human melanoma gene regulatory network, where the desired steady-state distribution corresponds to down-regulation of the WNT5A gene. The aim of perturbation in this case is to force the network away from its initial steady-state distribution associated with melanoma and into a benign state corresponding to a healthy cell. Steady-state distributions of gene regulatory networks have been associated with phenotypes such as cell proliferation and apoptosis.

Current biotechnology methods, however, may not easily translate into implementation of optimal control strategies. For example, in order to implement the proposed optimal control of the gene regulatory melanoma network in a laboratory setting, the probability that the network is in a given state should be determined experimentally by observing the dynamics of the melanoma cell over a period of time. We nonetheless believe that future translational research efforts will benefit from the proposed mathematically rigorous modeling and derivation methodology introduced in this paper provided it is properly adapted by incorporating clinical constraints dictated by the state-of-the-art molecular biotechnology techniques. For instance, although we have adopted the minimization of the energy, or the Frobenius norm, as the cost function used for determining the optimal perturbation control, other clinical objective criteria can and should be considered for determination of the optimal perturbation strategy in order to more closely adhere to clinical concerns. Furthermore, despite our focus on perturbation of the gene expression levels, other clinical control parameters can and should be considered in the future as potentially far more powerful methods for molecular network perturbation for gene regulation.

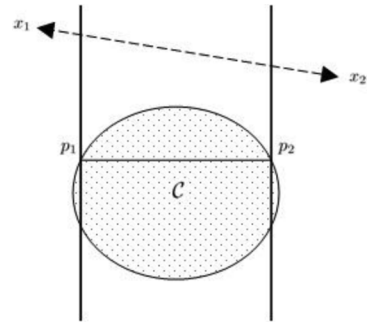

Fig. 5. Graphical illustration of the proof of Lemma 3.

Appendix A
Proof of Lemmas

Proof of Lemma 1: The proof follows immediately from the Jordan decomposition of the matrix $\boldsymbol{P}$. Since $\boldsymbol{P}$ is stochastic, 1 is the largest magnitude eigenvalue, and for any other eigenvalue λ, we have $|\lambda| \leq 1$. Therefore, $\boldsymbol{P}$ converges towards its steady-state distribution if and only if $\mathrm{SLEM}(\boldsymbol{P}) < 1$. ■

Proof of Lemma 3: Consider the line (p_1, p_2). Let $\mathcal{H}_1$ (resp., $\mathcal{H}_2$) be the hyperplane orthogonal to (p_1, p_2) at p_1 (resp., p_2). Then, x_1 must be to the left of $\mathcal{H}_1$ (see Fig. 5), otherwise some point strictly inside the segment $]p_1, p_2[\subseteq \Gamma$ will be closer to x_1 than p_1. Similarly, x_2 must be to the right of $\mathcal{H}_2$. Therefore, $\|x_1 - x_2\| \geq \|p_1 - p_2\|$. ■

Appendix B
Proof of Propositions

Proof of Proposition 1: We will distinguish between the matrix $\mathbf{P}$ and its corresponding operator $\tilde{\mathbf{P}}$. Let us find a finite dimensional operator $\tilde{\mathbf{P}}$ that satisfies:

$$(i)\ \tilde{\mathbf{P}}\mathbf{1} = \mathbf{1}, \quad (ii)\ \tilde{\mathbf{P}}^t \pi_d = \pi_d. \tag{28}$$

Consider the basis $\mathcal{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ defined as follows: $\mathbf{y}_1 = \pi_d$, $\mathbf{y}_2 = (\mathbf{1}, \pi_d)\, \perp\, \mathbf{1}$ and $\mathbf{y}_i = \mathbf{x}_i$, for $3 \leq i \leq n$. We know that the operator $\tilde{\mathbf{P}}$ can be written in the following form

$$\tilde{\mathbf{P}} = \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} < \bullet, \mathbf{y}_j > \mathbf{x}_k. \tag{29}$$

Applying the operator to the vector $\mathbf{x}_1$, we obtain

$$\tilde{\mathbf{P}}\mathbf{x}_1 = \sum_{k=1}^{n} \alpha_{1k}\mathbf{x}_k = \alpha_{11}\mathbf{x}_1 + \sum_{k=2}^{n} \alpha_{1k}\mathbf{x}_k. \tag{30}$$

Since $\mathbf{x}_1 = \mathbf{1}$ by construction, the condition $\tilde{\mathbf{P}}\mathbf{1} = \mathbf{1}$ is equivalent to $\alpha_{11} = 1$ and $\alpha_{1k} = 0$, for $k = 2, \ldots, n$. Similarly, the condition $\tilde{\mathbf{P}}^t \pi_d = \pi_d$ is equivalent to $\alpha_{11} = 1$ and $\alpha_{j1} = 0$, for

$j = 2, \ldots, n$. Finally, the operator $\tilde{\mathbf{P}}$ satisfying the two conditions in (28) can be written as:

$$\tilde{\mathbf{P}} = < \bullet, \mathbf{y}_1 > \mathbf{x}_1 + \sum_{j=2}^{n} \sum_{k=2}^{n} \alpha_{jk} < \bullet, \mathbf{y}_j > \mathbf{x}_k. \tag{31}$$

Applying the operator $\tilde{\mathbf{P}}$ in (31) to the basis $\mathcal{X}$, we obtain the desired block form representation $\mathbf{P}_{\mathcal{X}}$. The condition $\mathbf{P} \geq 0$ is satisfied if and only if $< \tilde{\mathbf{P}}\mathbf{e}_i, \mathbf{e}_j > \geq 0$ for all $1 \leq i, j \leq n$ or equivalently $\mathbf{A}\mathbf{P}_{\mathcal{X}}\mathbf{A}^{-1} \geq 0$, where $\{\mathbf{e}_i\}_{i=1}^{n}$ denotes the canonical basis, i.e., the standard basis for the Euclidean space $\mathbb{R}^n$.

The fact that the magnitude of the maximum eigenvalue of the $(n-1) \times (n-1)$ submatrix of $*$'s is strictly less than unity follows from the condition that $\mathrm{SLEM}(\boldsymbol{P}) < 1$. Given the block-diagonal structure of $\boldsymbol{P}$, $\mathrm{SLEM}(\boldsymbol{P})$ is equal to the maximum eigenvalue of the $(n-1) \times (n-1)$ submatrix. ■

Proof of Proposition 2: We first observe that any feasible perturbation matrix $\boldsymbol{C} \in \bar{\mathcal{F}}$ can be written as

$$\boldsymbol{C} = \boldsymbol{C}_0 + \boldsymbol{V}, \tag{32}$$

where $\boldsymbol{V} \in \Gamma$. Therefore, the optimization problem over $\bar{\mathcal{F}}$ is equivalent to minimizing the Frobenius norm of $(\boldsymbol{C}_0 + \boldsymbol{V})$ subject to $\boldsymbol{V} \in \Gamma$. It is straightforward that the solution to this problem is

$$\boldsymbol{V}^* = -\boldsymbol{C}_0^{\Gamma}. \tag{33}$$

Therefore, $\boldsymbol{C}^* = \boldsymbol{C}_0 - \boldsymbol{C}_0^{\Gamma}$. ■

Proof of Proposition 3: Because $\boldsymbol{P}_0 + \boldsymbol{C}^*$ is a stochastic matrix, we have $\mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}^*) \leq 1$. Therefore, $\boldsymbol{C}^* \in \bar{\mathcal{F}}$. If $\mathrm{SLEM}(\boldsymbol{P}_0 + \boldsymbol{C}^*) < 1$, then $\boldsymbol{C}^* \in \mathcal{F}$ and because $\mathcal{F} \subseteq \bar{\mathcal{F}}$, $\boldsymbol{C}^*$ is also the optimal solution of the optimization problem in (7). ■

Proof of Proposition 4: We will prove each of the three facts separately.

1) We need the following Lemma from [19]
   Lemma 4 [19] Let

   $$\boldsymbol{P}_n = \boldsymbol{P}_0 + \boldsymbol{C}_n = (1 - \epsilon_n)\boldsymbol{P}^* + \epsilon_n \mathbf{1}\boldsymbol{\pi}_d^t, \tag{34}$$

   where $\boldsymbol{P}^* = \boldsymbol{P}_0 + \boldsymbol{C}^*$ and $0 < \epsilon_n \leq 1$. Then, we have

   $$\mathrm{SLEM}\,(\boldsymbol{P}_n) = (1 - \epsilon_n)\,\mathrm{SLEM}\,(\boldsymbol{P}^*). \tag{35}$$

   Because $\boldsymbol{P}^*$ is a stochastic matrix, we have $\mathrm{SLEM}(\boldsymbol{P}^*) \leq 1$. Therefore, from Lemma 4, we obtain $\mathrm{SLEM}(\boldsymbol{P}_n) < 1$ for $0 < \epsilon_n \leq 1$. It is easy to check that $\boldsymbol{C}_n$ satsfies conditions (i)–(iii). Hence, $\boldsymbol{C}_n \in \mathcal{F}, \forall\, n \in \mathrm{IN}$.
2) The fact that $\boldsymbol{C}_n \to \boldsymbol{C}^*$ is obvious given that $\epsilon_n \to 0$.
3) We have $\boldsymbol{C}_n \in \mathcal{F} \subseteq \bar{\mathcal{F}}$, and $\boldsymbol{C}^*$ is the unique Frobenius norm minimizer over $\bar{\mathcal{F}}$. Hence, $\|\boldsymbol{C}_n\|_F \geq \|\boldsymbol{C}^*\|$. Moreover, $\boldsymbol{C}^* \notin \mathcal{F}$. In particular, $\boldsymbol{C}_n \neq \boldsymbol{C}^*$, $\forall\, n \in \mathrm{IN}$. Therefore, we have strict inequality $\|\boldsymbol{C}_n\|_F > \|\boldsymbol{C}^*\|_F$, $\forall\, n \in \mathrm{IN}$.

■

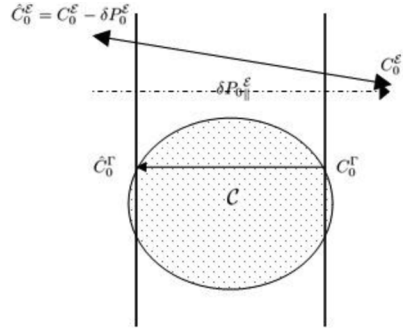

Fig. 6. Graphical illustration of the proof of Proposition 5.

Proof of Proposition 5: From proposition 2, we have

$$\boldsymbol{C}^* - \hat{\boldsymbol{C}}^* = (\boldsymbol{C}_0 - \boldsymbol{C}_0^{\Gamma}) - (\hat{\boldsymbol{C}}_0 - \hat{\boldsymbol{C}}_0^{\Gamma}) \tag{36}$$

$$= (\boldsymbol{C}_0 - \hat{\boldsymbol{C}}_0) + (\hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}) \tag{37}$$

$$= \delta\boldsymbol{P}_0 + (\hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}). \tag{38}$$

Let $\mathcal{E}$ be the subspace defined by the two equality constraints, i.e., $\mathcal{E} = \{\boldsymbol{V} : \boldsymbol{V}\mathbf{1} = \mathbf{0}, \boldsymbol{V}^t\boldsymbol{\pi}_d = \mathbf{0}\}$. Let us consider the decomposition of $\delta\boldsymbol{P}_0$ as

$$\delta\boldsymbol{P}_0 = \delta\boldsymbol{P}_0^{\mathcal{E}} + \delta\boldsymbol{P}_0^{\mathcal{E}^{\perp}}, \tag{39}$$

where $\delta\boldsymbol{P}_0^{\mathcal{E}}$ and $\delta\boldsymbol{P}_0^{\mathcal{E}^{\perp}}$ are, respectively, the projections of $\delta\boldsymbol{P}_0$ onto the subspace $\mathcal{E}$ and its orthogonal $\mathcal{E}^{\perp}$. Then (38) becomes

$$\boldsymbol{C}^* - \hat{\boldsymbol{C}}^* = \delta\boldsymbol{P}_0^{\mathcal{E}} + \delta\boldsymbol{P}_0^{\mathcal{E}^{\perp}} + (\hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}). \tag{40}$$

Let us further decompose $\delta\boldsymbol{P}_0^{\mathcal{E}}$ into the subspaces parallel and orthogonal to $\hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}$, i.e.,

$$\delta\boldsymbol{P}_0^{\mathcal{E}} = \delta\boldsymbol{P}_{0\|}^{\mathcal{E}} + \delta\boldsymbol{P}_{0\perp}^{\mathcal{E}}. \tag{41}$$

We have

$$\|\delta\boldsymbol{P}_0^{\mathcal{E}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2 = \|\delta\boldsymbol{P}_{0\|}^{\mathcal{E}} + \delta\boldsymbol{P}_{0\perp}^{\mathcal{E}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2$$
$$= \|\delta\boldsymbol{P}_{0\|}^{\mathcal{E}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2 + \|\delta\boldsymbol{P}_{0\perp}^{\mathcal{E}}\|^2.$$

From Lemma 3 and given that $\delta\boldsymbol{P}_{0\|}^{\mathcal{E}}$ is in the opposite direction as $\hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}$ (see Fig. 6), we have

$$\|\delta\boldsymbol{P}_{0\|}^{\mathcal{E}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2 \leq \|\delta\boldsymbol{P}_{0\|}^{\mathcal{E}}\|^2. \tag{42}$$

Therefore,

$$\|\delta\boldsymbol{P}_0^{\mathcal{E}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2 \leq \|\delta\boldsymbol{P}_{0\|}^{\mathcal{E}}\|^2 + \|\delta\boldsymbol{P}_{0\perp}^{\mathcal{E}}\|^2 = \|\delta\boldsymbol{P}_0{}^{\mathcal{E}}\|^2.$$

Finally, (40) becomes

$$\|\boldsymbol{C}^* - \hat{\boldsymbol{C}}^*\|^2 = \|\delta\boldsymbol{P}_0^{\mathcal{E}} + \delta\boldsymbol{P}_0^{\mathcal{E}^{\perp}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2 \tag{43}$$

$$= \|\delta\boldsymbol{P}_0^{\mathcal{E}} + \hat{\boldsymbol{C}}_0^{\Gamma} - \boldsymbol{C}_0^{\Gamma}\|^2 + \|\delta\boldsymbol{P}_0^{\mathcal{E}^{\perp}}\|^2 \tag{44}$$

$$\leq \|\delta\boldsymbol{P}_0^{\mathcal{E}}\|^2 + \|\delta\boldsymbol{P}_0^{\mathcal{E}^{\perp}}\|^2 = \|\delta\boldsymbol{P}_0\|^2. \tag{45}$$

■

ACKNOWLEDGMENT

The authors would like to extend their gratitude to Dr. R. Pal from Texas Tech University for providing the Human melanoma gene regulatory network dataset.

REFERENCES

[1] R. Demicheli and D. Coradini, "Gene regulatory networks: A new conceptual framework to analyse breast cancer behaviour," *Ann. Oncol.*, vol. 22, no. 6, pp. 1259–1265, 2011.
[2] I. Shmulevich, E. Dougherty, and W. Zhang, "From Boolean to probabilistic Boolean networks as models of genetic regulatory networks," *Proc. IEEE*, vol. 90, no. 11, pp. 1778–1792, Nov. 2002.
[3] N. Friedman, "Inferring cellular networks using probabilistic graphical models," *Science*, vol. 303, no. 5659, pp. 799–805, Feb. 2004.
[4] A. A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. D. Favera, and A. Califano, "ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context," *BMC Bioinform.*, vol. 7, no. 1, Mar. 2006.
[5] H. Hache, C. Wierling, H. Lehrach, and R. Herwig, "GeNGe: Systematic generation of gene regulatory networks," *Bioinformatics*, vol. 25, no. 9, pp. 1205–1207, 2009.
[6] D. Marbach, R. Prill, T. Schaffter, C. Mattiussi, D. Floreano, and G. Stolovitzky, "Revealing strengths and weaknesses of methods for gene network inference," *Proc. Nat. Acad. Sci.*, vol. 107, no. 14, pp. 6286–6291, Apr. 2010.
[7] M. Tan, M. Alshalalfa, R. Alhajj, and F. Polat, "Influence of prior knowledge in constraint-based learning of gene regulatory networks," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 8, no. 1, pp. 130–142, Feb. 2011.
[8] H. Fathallah-Shaykh, "Genomic discovery reveals a molecular system for resistance to ER and oxidative stress in cultured glioma," *Arch. Neurol.*, vol. 62, pp. 233–236, 2005.
[9] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. New York, NY, USA: Athena Scientific, 2007.
[10] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, pp. 167–173, May 2011.
[11] A. Datta, R. Pal, A. Choudhary, and E. Dougherty, "Control approaches for probabilistic gene regulatory networks—What approaches have been developed for addreassinig the issue of intervention?," *IEEE Signal Process. Mag.*, vol. 24, no. 1, pp. 54–63, Jan. 2007.
[12] B. Faryabi, G. Vahedi, J.-F. Chamberland, A. Datta, and E. R. Dougherty, "Optimal constrained stationary intervention in gene regulatory networks," *EURASIP J. Bioinform. Syst. Biol.*, vol. 2008, May 2008 [Online]. Available: http://bsb.eurasipjournals.com/content/2008/1/620767, Article ID 620767, doi: 10.1155/2008/620767
[13] X. Qian, I. Ivanov, N. Ghaffari, and E. R. Dougherty, "Intervention in gene regulatory networks via greedy control policies based on long-run behavior," *BMC Syst. Biol.*, vol. 3, no. 1, pp. 61–61, Jun. 2009.
[14] G. Vahedi, B. Faryabi, J. Chamberland, A. Datta, and E. Dougherty, "Intervention in gene regulatory networks via a stationary mean-first-passage-time control policy," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 10, pp. 2319–2331, Oct. 2008.
[15] 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, Jun. 1999.
[16] I. Shmulevich, E. R. Dougherty, and W. Zhang, "Control of stationary behavior in probabilistic Boolean networks by means of structural intervention," *J. Biolog. Syst.*, vol. 10, no. 4, pp. 431–445, 2002.
[17] Y. Xiao and E. Dougherty, "The impact of function perturbations in Boolean networks," *Bioinformatics*, vol. 23, no. 10, pp. 1265–1273, 2007.
[18] X. Qian and E. R. Dougherty, "Effect of function perturbation on the steady-state distribution of genetic regulatory networks: Optimal structural intervention," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 4966–4976, Oct. 2008.
[19] N. Bouaynaya, R. Shterenberg, and D. Schonfeld, "Inverse perturbation for optimal intervention in gene regulatory networks," *Bioinformatics*, vol. 27, no. 1, pp. 103–110, 2011.
[20] S. Kim, H. Li, E. R. Dougherty, N. Cao, Y. Chen, M. Bittner, and E. B. Suh, "Can Markov chain models mimic biological regulation?," *J. Biolog. Syst.*, vol. 10, no. 4, pp. 337–357, 2002.
[21] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. New York: Springer, 1983.
[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[23] M. Bittner, P. Meltzer, Y. Chen, Y. Jiang, E. Seftor, M. Hendrix, M. Radmacher, R. Simon, Z. Yakhini, A. Ben-Dor, N. Sampas, E. Dougherty, E. Wang, F. Marincola, C. Gooden, J. Lueders, A. Glatfelter, P. Pollock, J. Carpten, E. Gillanders, D. Leja, K. Dietrich, C. Beaudry, M. Berens, D. Alberts, and V. Sondak, "Molecular classification of cutaneous malignant melanoma by gene expression profiling," *Nature*, vol. 406, no. 6795, pp. 536–540, 2000.
[24] R. Pal, I. Ivanov, A. Datta, and E. R. Dougherty, "Generating Boolean networks with a prescribed attractor structure," *Bioinformatics*, vol. 21, pp. 4021–4025, 2005.

Nidhal Bouaynaya (M'09) received the B.S. degree in electrical engineering and computer science from the Ecole Nationale Supérieure de L'Electronique et de ses Applications (ENSEA), France, in 2002, the M.S. degree in electrical and computer engineering from the Illinois Institute of Technology, Chicago, in 2002, the Diplôme d'Etudes Approfondies in signal and image processing from ENSEA, France, in 2003, the M.S. degree in mathematics, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Chicago, in 2007.

In 2007, she joined the University of Arkansas at Little Rock, where she is currently an Assistant Professor in the Department of Systems Engineering. Her current research interests are in genomic signal processing, medical imaging, mathematical biology, and dynamical systems.

Dr. Bouaynaya won the Best Student Paper Award in Visual Communication and Image Processing 2006. She is currently serving as a Review Editor of *Frontiers in Systems Biology*.

Roman Shterenberg received the B.S. degree in physics in 1998, the M.S. and Ph.D. degrees in mathematics in 2000 and 2003, respectively, from St. Petersburg State University, Russia.

During 2005–2007, he was a Van Vleck Assistant Professor at University of Wisconsin-Madison. In 2007, he joined the University of Alabama at Birmingham where he is currently an Associate Professor in the Department of Mathematics. His research interests are in mathematical physics, spectral theory, inverse problems, and mathematical biology.

Dan Schonfeld (S'88–M'90–SM'05–F'12) received the B.S. degree in electrical engineering and computer science from the University of California at Berkeley, and the M.S. and Ph.D. degrees in electrical and computer engineering from The Johns Hopkins University, Baltimore, MD, in 1986, 1988, and 1990, respectively.

In 1990, he joined the University of Illinois at Chicago, where he is currently a Professor in the Departments of Electrical and Computer Engineering, Computer Science, and Bioengineering. He has authored more than 200 technical papers in various journals and conferences. His current research interests are in signal processing, image and video analysis, video retrieval and communications, multimedia systems, computer vision, medical imaging, and genomic signal processing.

Dr. Schonfeld has been elected University Scholar of the University of Illinois. He was a coauthor of a paper that won the Best Paper Award at the ACM Multimedia Workshop on Advanced Video Streaming Techniques for Peer-to-Peer Networks and Social Networking 2010. He was also the coauthor of papers that won the Best Student Paper Awards in Visual Communication and Image Processing 2006 and IEEE International Conference on Image Processing 2006 and 2007. He was elevated to the rank of Fellow of the IEEE "for contributions to image and video analysis." He was also elevated to the rank of Fellow of the SPIE "for specific achievements in morphological image processing and video analysis." He is currently serving as Deputy Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and Area Editor for Special Issues of the IEEE SIGNAL PROCESSING MAGAZINE.