

Intervention in gene regulatory networks with maximal phenotype alteration

Mohammadmahdi R. Yousefi¹ and Edward R. Dougherty^{1,2,*}¹Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 and²Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ 85004, USA

Associate Editor: Igor Jurisica

ABSTRACT

Motivation: A basic issue for translational genomics is to model gene interaction via gene regulatory networks (GRNs) and thereby provide an informatics environment to study the effects of intervention (say, via drugs) and to derive effective intervention strategies. Taking the view that the phenotype is characterized by the long-run behavior (steady-state distribution) of the network, we desire interventions to optimally move the probability mass from undesirable to desirable states. Heretofore, two external control approaches have been taken to shift the steady-state mass of a GRN: (i) use a user-defined cost function for which desirable shift of the steady-state mass is a by-product and (ii) use heuristics to design a greedy algorithm. Neither approach provides an optimal control policy relative to long-run behavior.

Results: We use a linear programming approach to optimally shift the steady-state mass from undesirable to desirable states, i.e. optimization is directly based on the amount of shift and therefore must outperform previously proposed methods. Moreover, the same basic linear programming structure is used for both unconstrained and constrained optimization, where in the latter case, constraints on the optimization limit the amount of mass that may be shifted to ‘ambiguous’ states, these being states that are not directly undesirable relative to the pathology of interest but which bear some perceived risk. We apply the method to probabilistic Boolean networks, but the theory applies to any Markovian GRN.

Availability: Supplementary materials, including the simulation results, MATLAB source code and description of suboptimal methods are available at <http://gsp.tamu.edu/Publications/supplementary/yousefi13b>.

Contact: edward@ece.tamu.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on February 20, 2013; revised on April 11, 2013; accepted on April 24, 2013

1 INTRODUCTION

Genetic regulatory networks (GRNs) refer to a class of models describing the multivariate functional relationships among a cohort of genes or their products. From a graphical perspective, genes are nodes in this network, and edges describe regulatory relationships between genes. These networks aim to model cellular control and how abnormal cell functions result from one or several breakdowns in the regulatory mechanisms. Thus, GRNs

are an essential part of translational medicine, whose ultimate goal is to develop therapies based on the disruption or mitigation of aberrant gene function contributing to the pathology of a disease (Dougherty *et al.*, 2010). Therapies usually involve some procedure and several drug candidates acting on various gene products with the aim of mitigating undesirable gene functions.

Developing therapeutic methods in the context of GRNs involves designing intervention strategies to alter the dynamics of the gene activity profiles (GAPs) of the network in some desired manner, thereby identifying potential drug targets (Datta and Dougherty, 2006; Dougherty and Datta, 2005; Dougherty *et al.*, 2010). Modeling GRNs via Markovian dynamical networks has received much attention because they capture uncertainty intrinsic to the interactions among genes or gene products at different levels. Furthermore, one can use the rich theory of Markov decision processes (MDPs) to formulate optimal intervention problems. In the present article, we choose probabilistic Boolean networks (PBNs) (Shmulevich *et al.*, 2002a) as our reference model for GRNs. The transition probabilities of PBNs are characterized by their associated Markov chains, and PBNs have played the major role in intervention studies.

Assuming that the collection of all possible GAPs constitutes the state space, we can partition it into *desirable* and *undesirable* cellular states according to the expression values of a given set of genes. An undesirable state can be associated with a phenotype representative of a cancerous state; e.g. cell growth in the absence of growth factors is undesirable. Although there might exist some states in the desirable set that do not represent healthy conditions, they can be classified as ‘desirable’ because they are not associated with the particular pathology defining the undesirable set (Qian and Dougherty, 2012). Once a model for the underlying GRN is assumed and desirable and undesirable sets are recognized, the objective is to find an optimal therapeutic strategy (control policy) with respect to a predefined objective function to drive the dynamics of the network from undesirable states to the desirable ones. This problem has been extensively studied within two basic intervention approaches in the context of PBNs, external control and structural intervention.

External control is generally based on flipping (or not flipping) the value of a specific gene (or possibly more than one), called the *control gene*. As the dynamic behavior of a PBN can be represented by a finite-state Markov chain, the first proposed intervention approach was to determine an optimal single-gene perturbation to the network based on mean first-passage times in Markov chains (Shmulevich *et al.*, 2002b). Attention, later,

*To whom correspondence should be addressed.

turned to dynamic programming-based finite-horizon (Datta *et al.*, 2003) and infinite-horizon external control in which the steady-state distribution (SSD) is altered (Pal *et al.*, 2006). These works were followed by several articles developing approaches to external intervention that take into account practical issues that arise from therapeutic constraints (Ching *et al.*, 2009; Yousefi *et al.*, 2012), biological complexity (Pal *et al.*, 2008) and computational limitations (Faryabi *et al.*, 2007; Ivanov *et al.*, 2010). Structural intervention, on the other hand, involves a one-time change in the network regulatory structure (wiring) so that the long-run behavior of the network is altered in a desired manner. Given a collection of potential structural changes, the problem is to find an optimal structural intervention resulting in a maximum alteration of the SSD toward the direction of desirable states and away from undesirable states (Qian and Dougherty, 2008; Shmulevich *et al.*, 2002c; Xiao and Dougherty, 2007).

In the main, optimal infinite-horizon intervention for GRNs has involved the specification of a cost function based on the current state of the system and the desirability of potential transitions. Thus, the long-run effect of the optimal policy on the network dynamics becomes a by-product of this cost function-based optimization problem. As phenotype is associated with steady-state behavior, as in the case of attractor cycles in BNs and PBNs (Kauffman, 1993), from a practical perspective, it would be better to determine optimality directly in terms of long-run behavior. Direct optimization has been addressed to some extent in Qian *et al.* (2009), where the authors propose three classes of greedy stationary intervention policies that bypass the need for a user-defined cost function and directly use long-run behavior as the optimization criterion to reduce the mass of the SSD corresponding to undesirable states and increase the mass corresponding to desirable states (Qian *et al.*, 2009).

In this article, we rigorously formulate this intervention problem and provide an optimal intervention policy with a computational complexity equivalent to solving a linear program (LP) optimization problem in Section 2.4 for a general cost function and in Section 2.5 for maximal phenotype alteration. We also consider a variant of this optimization problem where one might constrain the steady-state probability mass of some ‘ambiguous’ states in the network while finding the optimal intervention policy. This is especially important in therapeutic methods because it is prudent to avoid introducing new probability mass to states associated with unknown phenotypes (Qian and Dougherty, 2012). We demonstrate the performance of optimal policies using synthetically generated networks and two real networks derived from the metastatic melanoma and mammalian cell cycle.

2 SYSTEMS AND METHODS

2.1 PBNs

PBNs address uncertainty in the structure of BNs or the dynamics of state transitions (Shmulevich *et al.*, 2002a). We restrict ourselves to binary PBNs, it being assumed that the value of each gene is quantized to two levels 0 or 1, where 0 corresponds to an unexpressed (OFF) gene and 1 corresponds to an expressed (ON) gene. The mathematical theory extends directly to PBNs with discrete-valued nodes.

DEFINITION 1. A Boolean network $\text{BN}(V, F)$ is fully characterized by a set of n nodes, $V = \{v^1, v^2, \dots, v^n\}$, (representing genes $\{g^1, g^2, \dots, g^n\}$ or

their products) and a list of Boolean functions $F = \{f^1, f^2, \dots, f^n\}$ describing the functional relationships between the nodes. The Boolean function $f^i : \{0, 1\}^i \mapsto \{0, 1\}$ determines the value of node i at time $k + 1$, given the value of its predictors at time k by

$$v_{k+1}^i = f^i(v_k^1, v_k^2, \dots, v_k^{j_i}),$$

where $\{v_k^1, v_k^2, \dots, v_k^{j_i}\}$ corresponds to values of $\{g^1, g^2, \dots, g^{j_i}\}$ as j_i predictors of gene g^i .

In a BN, all genes are assumed to update synchronously in accordance with the Boolean functions assigned to them. A $\text{BN}(V, F)$ evolves according to a genome-wide GAP, $\mathbf{v}_k = (v_k^1, v_k^2, \dots, v_k^n)$ at time k , forming the state space of size 2^n . There is a bijection between \mathbf{v} and its decimal representation $x \in \bar{S} = \{0, 1, \dots, 2^n - 1\}$ by $x = \sum_{i=1}^n 2^{n-i} v^i$. In this context, genes that are reflective of an undesirable state are called *target genes*. \bar{S} is partitioned into subsets of *desirable* and *undesirable* states, denoted by \bar{D} and \bar{U} , respectively. Given an initial state, a BN will eventually enter a set of states, called an *attractor cycle*. The set of states leading to that attractor cycle is known as the *basin of attraction* (BOA) (Dougherty *et al.*, 2010).

DEFINITION 2. A $\text{PBN}(V, \mathbf{F}, P, q, p)$ is characterized by a set of n nodes, $V = \{v^1, v^2, \dots, v^n\}$, and a set of m constituent BNs, $\mathbf{F} = \{F^1, F^2, \dots, F^m\}$, called contexts, a selection probability vector $P = \{p^1, p^2, \dots, p^m\}$ over \mathbf{F} , a network switching probability q and a random gene perturbation probability p . Random switching and gene perturbation are also assumed to be mutually exclusive.

At any time point, with probability q , the network dynamics may switch from the current governing constituent BN to another according to the selection probability vector P . It is assumed that the probability of switching to a new constituent network is independent of the current network. We also allow that the current network may switch to itself when a switch is called for (Dougherty *et al.*, 2010). In addition, with probability p , the current state of each gene in the network can be randomly flipped. The PBN is said to be *context-sensitive* if $q < 1$, the interpretation being that there are latent variables outside the network whose changes cause the model network to behave stochastically (Brun *et al.*, 2005). If $q = 1$, the PBN is called *instantaneously random*, the interpretation being that the uncertainty in the PBN arises from uncertainty in model inference (Shmulevich *et al.*, 2002a). Averaging over the various contexts reduces the transition probability matrix (TPM) of a context-sensitive PBN to the instantaneously random PBN with identical parameters (Faryabi *et al.*, 2009). By definition, a PBN inherits the attractor structures from its context BNs without perturbation. With sufficiently small perturbation probability p , the long-run behavior of a PBN will reflect the attractor structures within the context BNs (Brun *et al.*, 2005). A Boolean network with perturbation, BN_p , is a PBN in which $m = 1$.

Transition rules of any PBN can be modeled by a homogeneous Markov chain, whose states of the TPM are the GAPs of the underlying regulatory network (Shmulevich *et al.*, 2002a). Let $\mathcal{S} = \{(x, y) : x \in \bar{S}, y \in \{1, 2, \dots, m\}\}$ denote the state space of the PBN. The sets \bar{D} and \bar{U} corresponding to the set of desirable, and undesirable states can be defined as $\bar{D} = \{(x, y) : x \in \bar{D}, y \in \{1, 2, \dots, m\}\}$ and $\bar{U} = \{(x, y) : x \in \bar{U}, y \in \{1, 2, \dots, m\}\}$. We denote by $\{\mathbf{Z}_k \in \mathcal{S}, k = 0, 1, \dots\}$ the stochastic process of the state of the PBN that has both the information about the current constituent BN and GAP of the underlying network. Originating from state $\mathbf{i} \in \mathcal{S}$, the successor state $\mathbf{j} \in \mathcal{S}$ is selected randomly according to the transition probability \mathcal{P} , with its ij th element defined by $p_{ij} \triangleq P(\mathbf{Z}_{k+1} = \mathbf{j} | \mathbf{Z}_k = \mathbf{i})$ for all $k = 0, 1, \dots$. The transition probabilities of this Markov chain can be calculated as explained in Dougherty *et al.* (2010). Owing to the random gene perturbation, the equivalent Markov chain is ergodic and has a unique invariant measure, π , equal to its limiting distribution.

We assume that the PBN admits an external control input A from a set of actions, \mathcal{A} , specifying the type of intervention on a set of control genes. For instance, $A = 0$ may indicate no-intervention, and $A = 1$ may indicate that the expression level of a single gene, g^c , $c \in \{1, 2, \dots, n\}$, is flipped. In this intervention scenario, the control action $A = 1$ at state (x, y) replaces the row corresponding to the state (x, y) in the original TPM of the underlying Markov chain by the row corresponding to the state (\tilde{x}, y) , where the binary representation of \tilde{x} is the same as x except in bit v^c , where it is flipped. Let $\{A_k \in \mathcal{A}, k = 0, 1, \dots\}$ denote the stochastic process of actions taken. The law of motion for the controlled network is represented by a matrix $\mathcal{P}(a)$ with its ij th element defined as

$$p_{ij}(a) = P(\mathbf{Z}_{k+1} = \mathbf{j} | \mathbf{Z}_k = \mathbf{i}, A_k = a), \quad (1)$$

this being the probability of going to \mathbf{j} at time $k+1$ starting from state \mathbf{i} and taking action a at time k . As the original Markov chain is ergodic, the controlled chain will also be ergodic having a unique invariant measure. The action process is stochastic in two ways: (i) the state process is stochastic, and the action process is a mapping from the entire history of states and actions to the action space so that the action process is also stochastic and (ii) we allow the actions to be random depending on the history and current state of the system. Indeed, the pair (\mathbf{Z}_k, A_k) is a stochastic process.

2.2 Constrained MDPs

Discrete time MDPs constitute a class of sequential decision making problems where the system of interest evolves stochastically at discrete time units. One can observe system states at times $k = 0, 1, \dots, N$, where N , called the 'horizon' may be finite or infinite. At each k , a decision maker can alter the costs or dynamics of the system by choosing some parameters, called 'actions'. The dynamical movement of the system from one state to another is completely characterized, given the current state and action taken at this state via a probability vector over all possible next states. The goal is to optimize an objective function (e.g. the infinite-horizon expected average cost) and find an optimal control policy inducing the optimal cost. More than one objective cost may exist, and the decision maker minimizes one of the objectives subject to constraints on the others. This class of MDPs are referred to as constrained MDPs (Altman, 1999) and mathematically defined as follows.

DEFINITION 3. A tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{P}(a), g, r\}$ defines a constrained MDP, where \mathcal{S} denotes the state space of a finite-state Markov chain, \mathcal{A} is a finite set of actions, $\mathcal{A}(\mathbf{i}) \subseteq \mathcal{A}$ denoting the set of actions available at state \mathbf{i} (we also denote by $\mathcal{K} = \{(\mathbf{i}, a) : \mathbf{i} \in \mathcal{S}, a \in \mathcal{A}(\mathbf{i})\}$ the set of state-action pairs), $\mathcal{P}(a)$ defined in Equation (1) is the TPM of the controlled Markov chain, $g : \mathcal{K} \rightarrow \mathbb{R}$ is an immediate cost, and $r : \mathcal{K} \rightarrow \mathbb{R}^D$ is a D -dimensional vector of immediate costs defining the D constraints.

Denote by $\{\mathbf{z}_k, k = 0, 1, \dots\}$ and $\{a_k, k = 0, 1, \dots\}$ the sequences of observed states and actions. A policy is a prescription for taking actions at each k . Actions may be taken in accordance with a random mechanism, possibly a function of the entire history of the system up to time k . To make this precise, let H_k be the random vector of previous states and actions occurring up to time k and h_k be the observed history, i.e. $h_k = (\mathbf{z}_0, a_0, \mathbf{z}_1, a_1, \dots, \mathbf{z}_k, a_k)$. A policy $\mu = (\mu_0, \mu_1, \dots, \mu_N)$ is a sequence prescribed by the decision maker that steers the dynamics of the underlying system. If the history h_{k-1} is observed up to time k , then the decision maker chooses an action $a \in \mathcal{A}(\mathbf{z}_k)$ with probability $\mu_k(a|h_{k-1}, \mathbf{z}_k)$, where

$$0 \leq \mu_k(a|h_{k-1}, \mathbf{z}_k) \leq 1, \quad \sum_{a \in \mathcal{A}(\mathbf{z}_k)} \mu_k(a|h_{k-1}, \mathbf{z}_k) = 1.$$

The class of all control policies μ is denoted by \mathcal{M} . The initial state \mathbf{i} of the Markov chain and any given policy μ determine a unique probability measure P_1^μ over the space of all trajectories of states and actions, which

correspondingly defines the stochastic processes \mathbf{Z}_k and A_k of the states and actions for the controlled system. E_1^μ denotes the corresponding expectation relative to which cost criteria are defined.

DEFINITION 4. The infinite-horizon expected average costs are

$$G(\mathbf{i}, \mu) = \limsup_{N \rightarrow \infty} \frac{1}{N+1} \sum_{k=0}^N E_1^\mu g(\mathbf{Z}_k, A_k),$$

$$R^d(\mathbf{i}, \mu) = \limsup_{N \rightarrow \infty} \frac{1}{N+1} \sum_{k=0}^N E_1^\mu r^d(\mathbf{Z}_k, A_k), \quad \text{for } d = 1, 2, \dots, D.$$

Given real numbers V^1, V^2, \dots, V^D , we can formulate the optimization problem in its most general form: for an initial state \mathbf{i} ,

$$\text{OP}_1 : \quad \min_{\mu \in \mathcal{M}} G(\mathbf{i}, \mu),$$

subject to $R^d(\mathbf{i}, \mu) \leq V^d$ for all $d = 1, 2, \dots, D$.

The set of all policies that satisfy the constraints is called the *feasible region*. Let $G^*(\mathbf{i})$ denote the optimal value achieved by the optimization procedure. A feasible policy μ^* is optimal if $G^*(\mathbf{i}) = G(\mathbf{i}, \mu^*)$, for $\mathbf{i} \in \mathcal{S}$. The search space \mathcal{M} is infinite (possibly uncountable); however, one may identify classes within \mathcal{M} that possess beneficial properties. We consider three classes: Markov policies, stationary policies and stationary deterministic policies, denoted by \mathcal{M}_M , \mathcal{M}_S and \mathcal{M}_D , respectively. \mathcal{M}_M includes policies for which μ_k is only a function of \mathbf{z}_k for any k . \mathcal{M}_S is a subset of \mathcal{M}_M and includes policies for which μ_k is time invariant (not dependent on k). This means that the probabilities that the control policy μ assigns to different actions only depend on the state and not the time. For example, for a PBN with transition probabilities defined in Equation (1) and any policy $\mu \in \mathcal{M}_S$, the underlying stochastic process becomes a stationary Markov chain with the set, $\mathcal{Q}(\mu)$, of transition probabilities defined by

$$q_{ij}(\mu) = \sum_{a \in \mathcal{A}(\mathbf{i})} p_{ij}(a)\mu(a|\mathbf{i}), \quad (2)$$

for all $\mathbf{i}, \mathbf{j} \in \mathcal{S}$. The third class, \mathcal{M}_D , is the set of all stationary deterministic policies such that $\mu(a|\mathbf{i})$ is either 0 or 1 for every $\mathbf{i} \in \mathcal{S}$. In this case, $\mu : \mathcal{S} \rightarrow \mathcal{A}$ is a single-valued transformation from the states to the actions.

Usually in the context of MDP problems, minimizing an objective function (the infinite-horizon expected total discounted cost or the expected average cost) with no constraints, i.e. $D=0$, is carried out by formulating a set of dynamic programming functional equations, known as the *Bellman optimality equations*. Using these equations, it can be shown that the optimal control policies belong to the class \mathcal{M}_D of policies. Hence, instead of searching the entire space of all history-dependent randomized control policies for the optimal solution, one can focus only on a set of coupled minimization problems over a (much smaller) set of actions.

2.3 Suboptimal phenotype alteration

Our goal is to find an intervention policy to maximally shift the long-run probability mass of undesirable states to desirable ones. Several algorithms have been proposed for intervention in Markovian GRNs that avoid using a user-defined cost function and work directly with the transition probabilities of the Markov chain associated with the network to approximate this goal. These algorithms are motivated by heuristics and obtain suboptimal policies. Owing to space limitations, here we only list the algorithms, with detailed descriptions being given in the Supplementary Material: (i) mean-first-passage-time (MFPT) control policy (Vahedi *et al.*, 2008), (ii) BOA control policy (Qian *et al.*, 2009), (iii) SSD control policy (Qian *et al.*, 2009) (iv) and conservative SSD (CSSD) control policy (Qian *et al.*, 2009).

Under the MFPT, BOA, SSD and CSSD control policies, as well as for user-defined cost functions, there might be an introduction of significant mass at states representing complications and harmful side effects to healthy cells, even though these are classified in the set of desirable states because they are not undesirable relative to the particular pathology of interest.

In Section 2.8.2, we will discuss a 10-gene network involving the gene WNT5A and explain why an intervention that downregulates WNT5A may have the beneficial effect of suppressing metastatic phenotypes. For this reason, networks involving WNT5A have been used in a number of control studies (Datta *et al.*, 2003; Faryabi *et al.*, 2009; Pal *et al.*, 2006). In particular, in Qian and Dougherty (2008), structural intervention consisting of a one-time change to the governing regulatory logic was applied to a seven-gene network containing WNT5A with the aim downregulating it in the long-run. Among the genes in the network, the minimum steady-state probability mass for upregulated WNT5A was obtained by perturbing the function determining the status of WNT5A. Unfortunately, this intervention had the effect of placing a large mass (> 0.4) at a state that had virtually no mass in the original SSD and in which STC2 is upregulated. Several studies have shown that STC2 plays a role in carcinogenesis. For instance, STC2 is upregulated in breast and ovarian cancer cells, following exposure to hypoxia (Law and Wong, 2010a), and high levels of STC2 expression are associated with increased invasiveness and metastasis in ovarian cancer cells (Law and Wong, 2010b). Hence, the unconstrained optimal structural intervention should be avoided. A better alternative would be to perturb the machinery governing RET1, where the results in Qian and Dougherty (2008) show that an appropriate perturbation of its regulatory logic reduces the steady-state mass of the upregulated WNT5A states almost as much as direct WNT5A intervention, and without the side effect of significantly upregulating STC2.

From a general perspective, it may be prudent to restrict the newly introduced mass to states associated with known healthy phenotypes. To achieve this end, the set \mathcal{D} of desirable states can be further partitioned into two categories, \mathcal{D}_h representing known healthy states (those known to be associated with a healthy phenotype) and \mathcal{D}_a representing ambiguous states that may (or certainly will) lead to phenotypes that, although not undesirable relative to the pathology of immediate interest, are themselves known to be undesirable or of which nothing is known. The goal is to find a control policy that shifts the long-run probability mass of undesirable states as much as possible and at the same time keeps the long-run probability mass of each ambiguous state below some level. Being that phenotypes are associated with attractors, a conservative approach is to define the ambiguous states as those states in \mathcal{D} that belong to the nonattractor states of the original network so that the control will not introduce new attractors (phenotypes). This approach is not only conservative but also does not require prior knowledge of which states in \mathcal{D} are, or are not, associated with pathological phenotypes. In the WNT5A example just cited, the intervention leading to the high-mass state in which STC2 was upregulated would not have been allowed under this criterion. The constrained SSD (conSSD) and constrained CSSD (conCSSD) control policies aim to find such (suboptimal) interventions (Qian and Dougherty, 2012).

None of the six long-run-based heuristic algorithms use a rigorous optimization to find optimal intervention policies with respect to the objectives and constraints.

2.4 Occupation measures and the primal LP

To proceed optimally in the case of constrained problems, we note that it has been shown that the objective functions for different cost criteria are a linear function of ‘occupation measures’ for stationary control policies. A salient consequence is that the original problem of finding the optimal cost and control policy can be transformed into an LP, referred to as the *primal LP*, where the optimization variables are the occupation measures (Derman, 1970; Kallenberg, 1983). The optimal solutions to this primal

LP determine, as a one-to-one relationship, the optimal stationary control policies (Altman, 1999). Let us first state an assumption that we will make throughout the rest of this article.

ASSUMPTION 1. The MDP is assumed to be unichain, meaning that for any $\mu \in \mathcal{M}_D$, the corresponding Markov chain described by the transition probabilities $\mathcal{Q}(\mu)$, defined in Equation (2), has at most one ergodic class and a (perhaps empty) set of transient states.

This assumption holds for the case of controlled PBNs, as the Markov chain corresponding to a given PBN is ergodic due to random gene perturbation. We now define the occupation measures.

DEFINITION 5. For any given initial state \mathbf{i} and policy μ , and any state-action pair \mathbf{j} and $a \in \mathcal{A}(\mathbf{j})$, the occupation measure is

$$v_{ja}^N(\mathbf{i}, \mu) = \frac{1}{N+1} \sum_{k=0}^N P_1^\mu(\mathbf{Z}_k = \mathbf{j}, A_k = a).$$

In other words, under any policy μ and given $\mathbf{Z}_0 = \mathbf{i}$, $v_{ja}^N(\mathbf{i}, \mu)$ is the expected frequency, up to time N , of entrances into state \mathbf{j} when action a is taken. Let $v^N(\mathbf{i}, \mu)$ denote the matrix of $v_{ja}^N(\mathbf{i}, \mu)$ over all \mathbf{j} and a . Let $\mathcal{V}(\mathbf{i}, \mu)$ be the nonempty set of all limit points of the sequence $\{v^N(\mathbf{i}, \mu), N = 0, 1, \dots\}$. Then, for any $\mathbf{i} \in \mathcal{S}$, $\mu \in \mathcal{M}$ and $v \in \mathcal{V}(\mathbf{i}, \mu)$,

$$\sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja} g(\mathbf{j}, a) \leq G(\mathbf{i}, \mu),$$

with equality holding for some $v \in \mathcal{V}(\mathbf{i}, \mu)$ —in particular, it holds when $\mu \in \mathcal{M}_S$ (Altman, 1999). Moreover, the union of all $\mathcal{V}(\mathbf{i}, \mu)$ over all $\mu \in \mathcal{M}$, where the $\mathcal{V}(\mathbf{i}, \mu)$ are singletons, is equal to the union of all $\mathcal{V}(\mathbf{i}, \mu)$ over all $\mu \in \mathcal{M}_S$ (Derman, 1970). If $\mu \in \mathcal{M}_S$, then the $\mathcal{V}(\mathbf{i}, \mu)$ are all singletons and independent of the initial state \mathbf{i} . Furthermore, for $v_{ja}(\mathbf{i}, \mu) \in \mathcal{V}(\mathbf{i}, \mu)$, the Markov chain with transition probabilities $\mathcal{Q}(\mu)$, defined in Equation (2), has a unique invariant probability measure vector $\pi(\mu)$, which satisfies

$$\pi_j(\mu) = \sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja}(\mathbf{i}, \mu), \quad \text{for all } \mathbf{j}, \mathbf{i} \in \mathcal{S}, \quad (3)$$

$$\pi(\mu) = \pi(\mu) \mathcal{Q}(\mu), \quad (4)$$

$$\sum_{\mathbf{j} \in \mathcal{S}} \pi_j(\mu) = 1, \quad \pi_j(\mu) \geq 0, \quad \text{for all } \mathbf{j} \in \mathcal{S}. \quad (5)$$

Therefore, one only needs to search the space of \mathcal{M}_S for the optimal solution by solving the following LP problem (Altman, 1999):

$$\min_v \sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja} g(\mathbf{j}, a),$$

LP₂ :

$$\text{subject to } \begin{cases} \sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja} d^d(\mathbf{j}, a) \leq V^d, d = 1, 2, \dots, D, \\ \sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja} = \sum_{\mathbf{i} \in \mathcal{S}} \sum_{a \in \mathcal{A}(\mathbf{i})} v_{ia} p_{ij}(a), \mathbf{j} \in \mathcal{S}, \\ \sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja} = 1, \\ v_{ja} \geq 0, \text{ for all } \mathbf{j} \in \mathcal{S}, a \in \mathcal{A}(\mathbf{j}). \end{cases}$$

Let v^* be a minimizing argument of the **LP₂** problem. Under Assumption 1, one can recover an optimal randomized stationary policy μ^* by

$$\mu^*(a|\mathbf{j}) = \frac{v_{ja}^*}{\sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja}^*}, \quad (6)$$

whenever $\sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja}^* > 0$, and if $\sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja}^* = 0$, then $\mu^*(a|\mathbf{j}) = 1$ for some $a \in \mathcal{A}(\mathbf{j})$. As the Markov chain for a controlled PBN is ergodic, there are no transient states. Hence, $\sum_{a \in \mathcal{A}(\mathbf{j})} v_{ja}^* \neq 0$.

THEOREM 1. (Altman, 1999) The optimization problems **OP₁** and **LP₂** are equivalent, and there is a one-to-one correspondence between their feasible (and optimal) solutions.

THEOREM 2. (Ross, 1989) If **LP₂** is feasible and μ^* is an optimal policy constructed by the aforementioned procedure, then there exists a list of at most $|\mathcal{S}| + D$ actions for μ^* so that μ^* requires randomization in at most D states.

If there are no constraints on the expected average cost criteria, i.e. $D=0$, then there is no randomization and $\mu^* \in \mathcal{M}_D$. The main result of this section can be summarized as follows: Under Assumption 1, for every policy in \mathcal{M} , there exists a stationary policy in \mathcal{M}_S that achieves the same limit point for the occupation measures; hence, we only need to consider the set \mathcal{M}_S for any optimization problem involving occupation measures of state and action in its objective and constraints.

2.5 Maximal phenotype alteration

We desire an intervention policy that maximally shifts the long-run probability mass of undesirable states to desirable states. This objective function essentially concerns the long-run behavior of the occupation measures marginalized over the actions. Thus, based on the discussion at the end of Section 2.4, we can limit the policy space to \mathcal{M}_S without loss of generality.

Let $\mathcal{A} = \mathcal{A}(\mathbf{j}) = \{0, 1\}$ for all $\mathbf{j} \in \mathcal{S}$. If policy $\mu \in \mathcal{M}_S$, then the amount of shift in the aggregated probability of undesirable states for a PBN controlled under μ is defined as

$$\Delta\pi_U(\mu) = \sum_{\mathbf{j} \in \mathcal{U}} \pi_{\mathbf{j}} - \sum_{\mathbf{j} \in \mathcal{U}} \pi_{\mathbf{j}}(\mu), \quad (7)$$

where π and $\pi(\mu)$ are the unique vectors of the invariant probability measure for the Markov chains governed under the TPMs \mathcal{P} and $\mathcal{Q}(\mu)$, respectively, which also satisfy Equations (3) to (5). In general, $-1 \leq \Delta\pi_U(\mu) \leq 1$ and our goal is to maximize it.

LEMMA 1. $\Delta\pi_U$ is maximized by solving the **LP₂** problem with the immediate cost function g being 1 for undesirable states and 0 otherwise.

PROOF. It can be easily seen from Equation (7) that maximizing $\Delta\pi_U(\mu)$ is equivalent to minimizing $\sum_{\mathbf{j} \in \mathcal{U}} \pi_{\mathbf{j}}(\mu)$, as π is fixed. Also, if we let the immediate cost function $g(\mathbf{Z}_k, A_k)$ take the form:

$$g(\mathbf{j}, a) = \begin{cases} 1, & \text{if } \mathbf{j} \in \mathcal{U}, \\ 0, & \text{otherwise,} \end{cases}$$

we have

$$\sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a}(\mathbf{i}, \mu) g(\mathbf{j}, a) = \sum_{\mathbf{j} \in \mathcal{U}} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a}(\mathbf{i}, \mu) = \sum_{\mathbf{j} \in \mathcal{U}} \pi_{\mathbf{j}}(\mu),$$

for all \mathbf{i} and any $\mu \in \mathcal{M}_S$, where we used Equation (3) for the second equality. Therefore, it can be verified that we need to again solve **LP₂**, with the suggested choice of g , to find the optimal v , which at the same time maximizes the shift $\Delta\pi_U$.

2.5.1 Unconstrained optimal intervention If there are no constraints on the cost criteria ($D=0$) for the shift maximization problem, then we can rewrite **LP₂** as

$$\mathbf{LP}_3 : \begin{cases} \min_v \sum_{\mathbf{j} \in \mathcal{U}} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a}, \\ \text{subject to} \begin{cases} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a} = \sum_{\mathbf{i} \in \mathcal{S}} \sum_{a \in \mathcal{A}} v_{\mathbf{i}a} p_{\mathbf{i}\mathbf{j}}(a), \mathbf{j} \in \mathcal{S}, \\ \sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a} = 1, \\ v_{\mathbf{j}a} \geq 0, \text{ for all } \mathbf{j} \in \mathcal{S}, a \in \mathcal{A}. \end{cases} \end{cases}$$

The unconstrained (UC) optimal intervention policy μ_{uc}^* can be constructed using Equation (6) when v^* yields the minimum in **LP₃**. As there are no constraints, $\mu_{uc}^* \in \mathcal{M}_D$ so that μ_{uc}^* is stationary and deterministic.

2.5.2 Phenotypically constrained optimal intervention The **LP₃** problem aims to locate an optimal control policy inducing maximal shift from the undesirable mass to eradicate phenotypes associated with the pathology of interest. However, when there are ambiguous states in the network, the optimization problem can be stated as locating a control policy that maximizes the shift of undesirable mass while satisfying an upper bound constraint for newly introduced steady-state mass into ambiguous states in \mathcal{D}_a . Similar to the choice of g in the proof of Lemma 1, one can define r^d for $d = 1, 2, \dots, |\mathcal{D}_a|$, each corresponding to a state $\mathbf{j}^d \in \mathcal{D}_a$. The phenotypically constrained (PC) optimization problem can be formulated as

$$\mathbf{LP}_4 : \begin{cases} \min_v \sum_{\mathbf{j} \in \mathcal{U}} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a}, \\ \text{subject to} \begin{cases} \sum_{a \in \mathcal{A}} v_{\mathbf{j}^d a} \leq V^d, d = 1, 2, \dots, |\mathcal{D}_a|, \mathbf{j}^d \in \mathcal{D}_a, \\ \sum_{a \in \mathcal{A}} v_{\mathbf{j}a} = \sum_{\mathbf{i} \in \mathcal{S}} \sum_{a \in \mathcal{A}} v_{\mathbf{i}a} p_{\mathbf{i}\mathbf{j}}(a), \mathbf{j} \in \mathcal{S}, \\ \sum_{\mathbf{j} \in \mathcal{S}} \sum_{a \in \mathcal{A}} v_{\mathbf{j}a} = 1, \\ v_{\mathbf{j}a} \geq 0, \text{ for all } \mathbf{j} \in \mathcal{S}, a \in \mathcal{A}. \end{cases} \end{cases}$$

One can assume that $V^d = \lambda$ for all d . The PC optimal intervention policy μ_{pc}^* can be constructed using Equation (6) when v^* yields the minimum in **LP₄**. As there are $|\mathcal{D}_a|$ additional constraints on the optimization problem, $\mu_{pc}^* \in \mathcal{M}_S$ and it requires randomization in at most $|\mathcal{D}_a|$ states.

2.6 Computational complexity

In any optimization problem, computational complexity is a concern. In particular, how well does the algorithm scale to larger networks? In our case, there are $|\mathcal{S}| \times |\mathcal{A}|$ decision variables and $|\mathcal{S}|$ equality constraints in both **LP₃** and **LP₄** problems. **LP₄** has $|\mathcal{D}_a|$ additional inequality constraints. We use IBM ILOG CPLEX Optimizer using a dual simplex method to solve both **LP₃** and **LP₄** problems. Although, the worst-case computational complexity of the simplex method has been shown to be exponential in the number of decision variables and constraints, probabilistic analyses of the simplex method have indicated that the average complexity of this method is polynomial. Moreover, it has been proven that solving an LP, in general, requires a number of operations polynomial in the number of decision variables and constraints. Specifically, Karmarkar's algorithm, which belongs to the general class of interior point methods, solves LPs in polynomial time with reasonable efficiency (Megiddo, 1987). The computational complexities of all the suboptimal policies mentioned in the article are also reported to be either linear or polynomial.

Besides the guaranteed optimality of our method over the greedy algorithms, the linear programming approach is particularly useful when optimization problems involving nonlinear functions or side constraints are considered (Derman, 1970). In such cases, using a dynamic programming approach makes the analysis much more complicated and computationally prohibitive. Based on our estimates and the reported size of the linear programming problems that have been solved using IBM ILOG CPLEX Optimizer, it is possible with the current technology to find the optimal UC and PC policies for networks with 20 genes. Hence, it avoids the kinds of model reduction methodologies that have been developed in the context of GRN control (Ghaffari *et al.*, 2010; Ivanov *et al.*, 2007, 2010; Qian *et al.*, 2010), thereby avoiding the loss of optimality engendered by reductions.

2.7 Robustness

As with any optimization procedure, the algorithm is optimal so long as the assumed model is correct. In the case of GRNs, model fidelity is affected by several factors: data extraction, discretization, gene selection and network generation. Hence, robustness becomes an issue. Robustness refers to an algorithm's performance on models that are close to, but not equal to, the model on which it has been derived. The robustness of stationary control for PBNs has been considered in Pal *et al.* (2008), and a robust intervention strategy has been obtained by minimizing the worst-case cost over an uncertainty class of networks. Such a minimax control approach is typically conservative because it gives equal importance to extreme cases. Thus, a Bayesian approach was formulated in Pal *et al.* (2009). A Bayesian approach requires a prior distribution on the uncertainty class of networks and therefore is dependent on significant prior knowledge, which may not be available. Here, we describe the perturbation bounds discussed in Pal *et al.* (2008), where uncertainty is studied in the context of the transition probabilities. The intervention strategy is derived on the estimated TPM, \mathcal{P} , and is applied to the actual network TPM, $\tilde{\mathcal{P}}$. The basic question is that, for a given control policy, how does the mismatch between \mathcal{P} and $\tilde{\mathcal{P}}$ affect the SSD of the controlled network? Although the discussion in Pal *et al.* (2008) applies to stationary deterministic control and the equivalent and our general setting is of wider scope, both the UC and PC algorithms have stationary control so that the analysis of Pal *et al.* (2008) is applicable.

Given that the class of allowed interventions consists of flipping the value of a gene, application of a stationary policy μ derived from the uncontrolled estimated TPM \mathcal{P} converts \mathcal{P} to a controlled TPM $\mathcal{Q}(\mu)$, where $\mathcal{Q}(\mu) = T(\mu)\mathcal{P}$ and $T(\mu)$ represents a stochastic matrix, which has at most $|\mathcal{A}|$ nonzero entries adding up to 1 in each row (these nonzero entries are found from the control policy μ). Let π and $\pi(\mu)$ denote the SSDs corresponding to \mathcal{P} and $\mathcal{Q}(\mu)$, respectively. With $\tilde{\mathcal{P}}$ being the actual uncontrolled TPM, $\tilde{\mathcal{Q}}(\mu) = T(\mu)\tilde{\mathcal{P}}$ is the controlled TPM resulting from applying $T(\mu)$ to $\tilde{\mathcal{P}}$. Let $\tilde{\pi}$ and $\tilde{\pi}(\mu)$ denote the SSDs of $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{Q}}(\mu)$, respectively. Robustness concerns the difference $\tilde{\pi}(\mu) - \pi(\mu)$ based on the estimation error $E \triangleq \mathcal{P} - \tilde{\mathcal{P}}$.

Assuming that the actual and estimated networks possess the same state space, which allows for different topologies and different regulatory functions, the SSD difference can be bounded by $|\pi - \tilde{\pi}|_q \leq K\|E\|_\infty$ where $q = 1$ or ∞ and $K > 0$ are some constants. K is called a *condition number*. Some condition numbers will yield tighter bounds than the others (Cho and Meyer, 2001). Pal *et al.* (2008) considered the *ergodicity coefficient*

$$\tau_1(\mathcal{P}) = \sup_{\substack{|x^T|_1=1 \\ x^T \mathbf{1}=0}} |x^T \mathcal{P}|_1,$$

where $\mathbf{1}$ denotes a column vector of appropriate length having all entries equal to 1 (Seneta, 1988). If $\tau_1(\mathcal{P}) \neq 1$, then (Pal *et al.*, 2008)

$$|\pi(\mu) - \tilde{\pi}(\mu)|_1 \leq \frac{1}{1 - \tau_1(\mathcal{P})} \|E\|_\infty.$$

Hence if, the error is small, then the SSD of the controlled actual process is close to the SSD of controlled estimated process from which the control policy has been derived, the bound depending on the ergodicity coefficient. Other perturbation bounds using different condition numbers were examined via simulations in Pal *et al.* (2008).

Having discussed the general case of Markov chains, Pal *et al.* (2008) went on to relate the general results to PBNs by showing how different classes of possible uncertainties for a PBN would translate into uncertainties in the TPM for the corresponding Markov chain.

2.8 Simulation setup

We design simulation studies on synthetically generated and two real networks. We find optimal and suboptimal control policies for both

UC and constrained problems and calculate their performance under different parameter choices. To reduce the complexity of graphs and tables, we do not consider the MFPT and BOA policies. From the perspective of demonstrating the optimality of the UC and PC algorithms, there is no loss in this omission, as it has been amply demonstrated that the SSD and CSSD policies generally outperform the MFPT and BOA policies relative to the criterion of shifting the SSD (Qian *et al.*, 2009). For constrained optimization problems, the set of ambiguous states, \mathcal{D}_a , and an upper-bound λ must be defined. In practice, these should be defined based on biomedical knowledge related to the GRN and treatment objectives. Here, we take an approach similar to Qian and Dougherty (2012). For PBNs, we define $\mathcal{D}_a = \{\mathbf{i} \in \mathcal{D} : \pi_{\mathbf{i}} \leq \tau\}$, where π is the invariant measure of the TPM of the uncontrolled network and τ is a user-defined threshold. We let $\tau = 1/2^n$ and set $\lambda = \max_{\mathbf{i} \in \mathcal{D}_a} \pi_{\mathbf{i}}$. For BNps, ambiguous states are the states in \mathcal{D} that belong to the nonattractor states of the original BNs so that the control will not introduce new attractors.

2.8.1 Synthetic networks We generate 2500 random PBNs and 2500 random BNps with different parameters and report the average performance across all networks as well as some statistics on the performance of each intervention policy. The performance of optimal policies might not substantially exceed that of suboptimal policies when averaged across randomly generated networks for a couple of reasons. First, randomly generated networks may have certain structures making them unresponsive to the intervention policies. Second, many networks might possess structure for which the suboptimal and optimal policies are almost identical. Be that as it may, the key point is that there are networks for which the optimal policy significantly outperforms the suboptimal ones. Here, we use the difference in the amounts of shifts made by optimal and suboptimal policies to quantify the gain made by using the optimal policy:

$$\Delta(\mu) = \Delta\pi_{\mathcal{U}}(\mu^*) - \Delta\pi_{\mathcal{U}}(\mu) = \sum_{j \in \mathcal{U}} \pi_j(\mu) - \sum_{j \in \mathcal{U}} \pi_j(\mu^*),$$

where $\pi(\mu^*)$ and $\pi(\mu)$ denote the invariant probability measures for the network under the optimal and suboptimal policies, respectively. Clearly, $0 \leq \Delta(\mu) \leq 1$. For each network $\Delta(\mu)$ is calculated for all previously discussed suboptimal policies, for the UC and constrained problems separately. As the networks are randomly generated, $\Delta(\mu)$ is a random variable. To show the effectiveness of optimal policies, we graph the tail probabilities, i.e. the empirical complementary cumulative distribution function (CCDF), of $\Delta(\mu)$.

To keep the computational time tractable, we consider instantaneously random PBNs, the states representing the GAPs at any given time, and $n = 6, 7$ or 8 genes. The state space is $\mathcal{S} = \{0, 1, 2, \dots, 2^n - 1\}$ and is generated from four equally likely constituent BNs with the maximum number of predictors ($j_i \leq \kappa$ for all i) for each Boolean function set to two or three. The *bias* of each PBN is the probability that each Boolean regulatory function takes on the value 1. We assume that it is taken randomly from a beta distribution with mean $\theta \in \{0.25, 0.5, 0.75\}$ and the standard deviation 0.01, thereby affecting the dynamics of a randomly generated BN. The gene perturbation probability is either 0.001 or 0.01. A similar setting is used for generating BNps, except that there is only one constituent BN. For any given BNp and PBN, the TPMs of the corresponding Markov chains are computed as explained in Section 2.1. We choose the control and target genes to be the least and most significant bits in the binary representation of states, respectively, and assume that downregulation of the target gene is undesirable.

The actual construction is done in the following manner. For each constituent BN in a random PBN, we first randomly select the predictors of each gene with all genes having the same probability of being selected. The value of each gene given the values of its predictors is now determined by a Bernoulli random variable whose probability of being 1 equals the bias (the bias is also randomly generated from a beta distribution with a mean θ and a given standard deviation). This process will construct a random truth table corresponding to a BN. Having all the

constituent BNs and gene perturbation probability defined, we directly calculate the TPM of an instantaneously random PBN without actually constructing it.

2.8.2 Metastatic melanoma network A metastatic melanoma network is derived from gene expression data collected in a study of metastatic melanoma (Bittner *et al.*, 2000; Dissanayake *et al.*, 2008; Weeraratna *et al.*, 2002). By manipulating the concentration level of WNT5A protein secreted by a melanoma cell line, one can directly affect the metastatic status of the cell as measured by the standard *in vitro* assays for metastasis. For intervention purposes, antibodies can be designed to bind with WNT5A and block it from activating its receptor. This will significantly weaken its ability to induce a metastatic phenotype, which suggests that reducing the WNT5A gene's action could reduce the chance of a melanoma metastasizing. Therefore, it is desirable to downregulate WNT5A as much as possible through appropriate intervention mechanisms (Datta *et al.*, 2003; Qian and Dougherty, 2008). We use a network composed of 10 genes from a set of 587 genes (Qian and Dougherty, 2012). The genes, listed from the most significant bit to the least significant bit, are WNT5A, pirin, S100P, RET1, MMP3, PHOC, MART1, HADHB, synuclein and STC2. This ordering of genes is only for demonstration purposes and does not affect our analysis. The regulatory relationships between these genes are presented in Table 1. We construct a BNp with $p=0.001$ and assume that the upregulation of WNT5A is undesirable. For the control genes, we choose each gene in the network as the control gene and find the optimal and suboptimal intervention policies in addition to their effects on the steady-state shift.

2.8.3 Mammalian cellcycle network To characterize the dynamical behavior of normal mammalian cells during the cell cycle, Faure *et al.* (2006) proposed a Boolean-logic regulatory network containing three key genes: Cyclin D (CycD), retinoblastoma (Rb) and p27. For a normal mammalian organism, cell division is coordinated with overall growth via extracellular signals controlling the activation of CycD. These signals indicate whether a cell should undergo cell division or remain in a resting state. Rb is a tumor-suppressor gene and is expressed in the absence of the cyclins that inhibit Rb by phosphorylation. Gene p27 is also active in the absence of the cyclins. An active p27 blocks the action of CycE or CycA and, hence, Rb can also be expressed, even in the presence of CycE or CycA, resulting in a stop in the cell cycle. In the wild-type cell-cycle network, when p27 is active, the cell cycle can be stopped. Following a proposed mutation for this network, we assume that p27 can never be activated (always OFF), thereby creating a situation where both CycD and Rb might be inactive (Faure *et al.*, 2006). Under these conditions, the cell can cycle in the absence of any growth factor, thereby causing undesirable proliferation. The mutated network has nine genes, CycD, Rb, E2F, CycE, CycA, Cdc20, Cdh1, UbcH10 and CycB, ordered from the most significant bit to the least significant bit in the binary representation. Similar to the metastatic melanoma network, this ordering is only for the sake of presentation. Table 2 lists the regulatory functions of the mutated cell-cycle network following Boolean logic.

We construct an instantaneously random PBN for this mutated network, where depending on the state of the extracellular signal that determines the state of CycD as being ON or OFF, there are two constituent BNs, which we assume to be equally likely. We set $p=0.001$. As cell growth in the absence of growth factors is undesirable, the undesirable states are those for which CycD and Rb are both downregulated (OFF). We consider all genes except CycD and Rb as the possible control gene and find the optimal and suboptimal intervention policies in addition to their effects on the steady-state shift.

Table 1. Regulatory functions of a metastatic melanoma network

Gene	Node	Predictor functions
WNT5A	v_1	$(v_3 \wedge v_5 \wedge \bar{v}_6) \vee (\bar{v}_5 \wedge v_6)$
pirin	v_2	$(\bar{v}_1 \wedge \bar{v}_3 \wedge v_5) \vee (v_1 \wedge \bar{v}_3 \wedge \bar{v}_5)$
S100P	v_3	v_7
RET1	v_4	$(\bar{v}_1 \wedge v_2 \wedge v_4) \vee (\bar{v}_2 \wedge v_4)$
MMP3	v_5	$(v_4 \wedge v_9) \vee (\bar{v}_9)$
PHOC	v_6	$(\bar{v}_4 \wedge \bar{v}_7) \vee (v_4 \wedge v_7 \wedge v_{10})$
MART1	v_7	v_7
HADHB	v_8	$(v_1 \wedge v_5) \vee (\bar{v}_5 \wedge \bar{v}_9) \vee (v_1 \wedge \bar{v}_5 \wedge v_9)$
synuclein	v_9	$(\bar{v}_1 \wedge \bar{v}_7 \wedge \bar{v}_{10}) \vee (v_4 \wedge \bar{v}_7 \wedge v_{10}) \vee v_7$
STC2	v_{10}	\bar{v}_3

Table 2. Regulatory functions of a mutated mammalian cell-cycle network

Gene	Node	Predictor functions
CycD	v_1	Extracellular signal
Rb	v_2	$(\bar{v}_1 \wedge \bar{v}_4 \wedge \bar{v}_5 \wedge \bar{v}_9)$
E2F	v_3	$(\bar{v}_2 \wedge \bar{v}_5 \wedge \bar{v}_9)$
CycE	v_4	$(v_3 \wedge \bar{v}_2)$
CycA	v_5	$(v_3 \wedge \bar{v}_2 \wedge \bar{v}_6 \wedge (\bar{v}_7 \wedge v_8))$ $\vee (v_5 \wedge \bar{v}_2 \wedge \bar{v}_6 \wedge (\bar{v}_7 \wedge v_8))$
Cdc20	v_6	v_9
Cdh1	v_7	$(\bar{v}_5 \wedge \bar{v}_9) \vee v_6$
UbcH10	v_8	$\bar{v}_7 \vee (v_7 \wedge v_8 \wedge (v_6 \vee v_5 \vee v_9))$
CycB	v_9	$(\bar{v}_6 \wedge \bar{v}_7)$

3 RESULTS AND DISCUSSION

The entire set of simulation results can be found in the Supplementary Material. Here, we provide some results that represent the general trends observed in the simulations.

3.1 Synthetic networks

We present results for the synthetic networks with seven genes. Corresponding results for six and eight genes are given in the Supplementary Material, where it is evident that the same trends are observed. Figure 1 shows the shift in the probability mass of undesirable states induced by the optimal and suboptimal policies, Equation (7), averaged over all random networks with various structural properties and for different values of θ . These graphs clearly show the optimality of μ_{uc}^* and μ_{pc}^* compared with their suboptimal counterparts. In general, the suboptimal policy μ_{cssd} yields the best results, close to optimal, among the suboptimal policies. However, $\mu_{concssd}$ does not induce as much shift as μ_{pc}^* , which is due to the fact that the search space of conCSSD policy is \mathcal{M}_D , whereas it is \mathcal{M}_S in the \mathbf{LP}_4 problem. The amount of average shift induced by μ_{pc}^* is significantly less than μ_{uc}^* owing to the additional constraints imposed on the optimization \mathbf{LP}_4 problem.

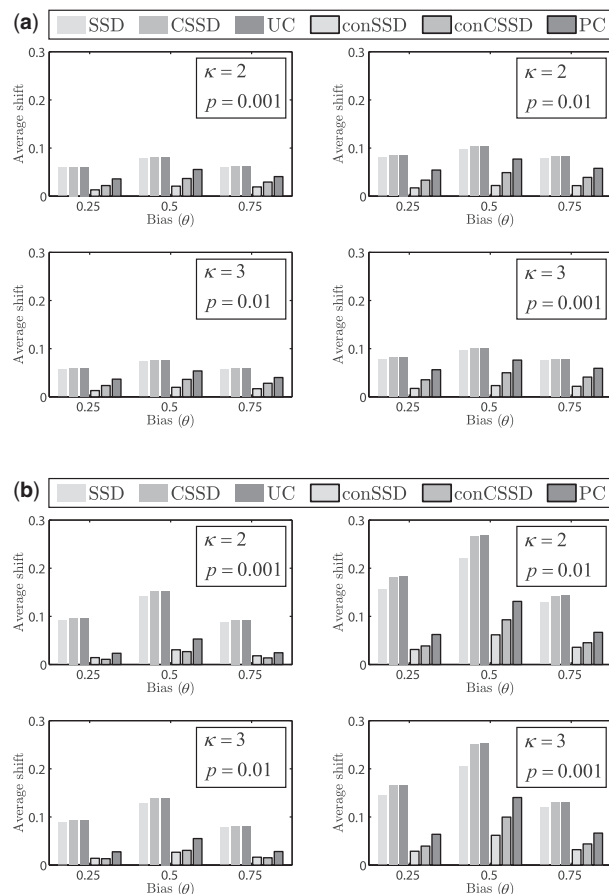


Fig. 1. Shift in the steady-state mass of undesirable states averaged across randomly generated networks with seven genes, various κ and gene perturbation probabilities p : (a) PBN; (b) BNP

Figure 2 shows the empirical CCDF of $\Delta(\mu)$ for different suboptimal intervention policies μ , and for the UC and PC problems separately, across 2500 random PBNs with seven genes, $\kappa = 3$ and $p = 0.001$. The graphs on the left column indicate that UC optimal and CSSD policies induce similar amounts of shift with a high probability. However, for the constrained problem, it is clear that PC optimal policies outperform conCSSD policies with a (realistically) significant probability. For example, for optimal policy μ_{pc}^* , with probability 0.1, we achieve an improvement of $> 5\%$ over $\mu_{concssd}$ suboptimal policy.

The results are somewhat different for BNPs. Figure 3 demonstrates the empirical CCDF of $\Delta(\mu)$ for different suboptimal intervention policies μ , corresponding to the UC and PC problems, across 2500 randomly generated BNPs with seven genes, $\kappa = 3$ and $p = 0.001$. With high probability, higher than those for PBNs, we achieve better performance by using optimal policies. For instance when $\theta = 0.5$, this figure shows that the optimal policy achieves an improvement of $> 5\%$ over $\mu_{concssd}$ with probability 0.2.

3.2 Real networks

As there is only one network, we cannot report the average shift or CCDF of $\Delta(\mu)$ for different suboptimal intervention policies.

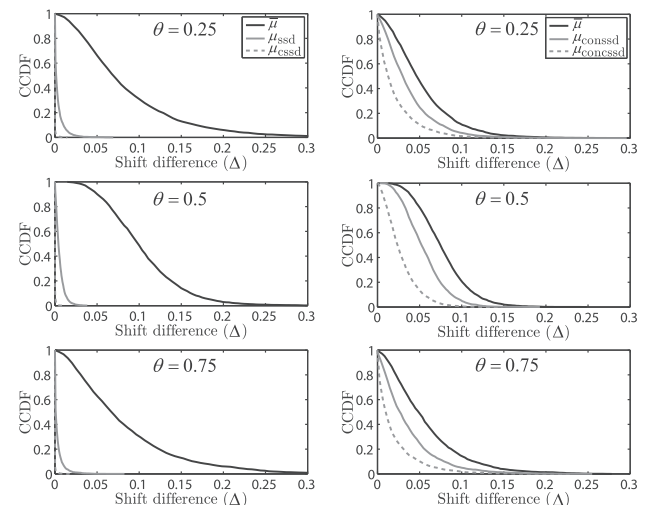


Fig. 2. Empirical CCDF of Δ for different intervention policies across 2500 random PBNs with seven genes, $\kappa = 3$ and $p = 0.001$. The left and right columns correspond to the UC and PC optimal policies, respectively

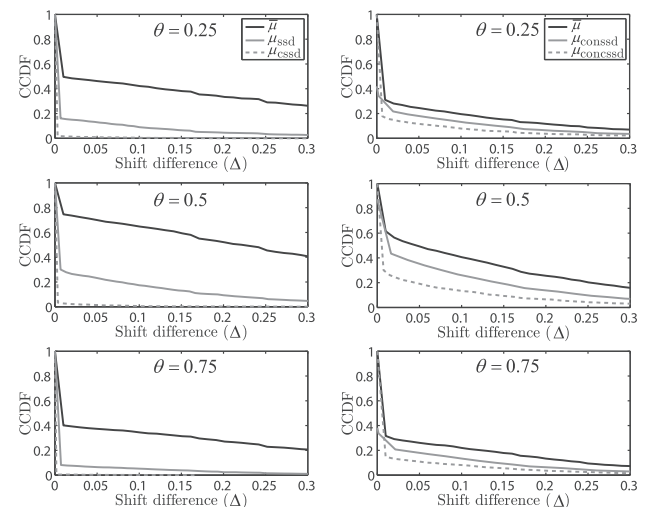


Fig. 3. Empirical CCDF of Δ for different intervention policies across 2500 random BNPs with seven genes, $\kappa = 3$ and $p = 0.001$. The left and right columns correspond to the UC and PC optimal policies, respectively

Table 3 shows the shifts made by all control policies for different potential control genes for the metastatic melanoma network. The aggregated probability mass of undesirable states for the uncontrolled network is $\sum_{j \in \mathcal{U}} \pi_j = 0.1711$. The maximum shift induced by the UC optimal policy occurs when either MMP3 or PHOC are the control genes. However, the maximum shift for the PC optimal policy occurs when MART1 is the control gene, which is an immunogen that the body sometimes recognizes and attacks via the immune system (Dissanayake *et al.*, 2008; Qian and Dougherty, 2012). A good amount of shift is achieved for both constrained and UC problems with MART1. Thus, MART1 can be viewed as a potential control gene for intervention. Overall, however, the results indicate that the best control gene depends on ones objectives and how much risk one is

Table 3. Shift in the steady-state mass of undesirable states of the metastatic melanoma network for different control genes and different control policies

Control	WNT5A	pirin	S100P	RET1	MMP3	PHOC	MART1	HADHB	synuclein	STC2
SSD	0.0632	0.1630	0.1679	0.1646	0.1701	0.1701	0.1684	0.0000	0.0029	0.1674
CSSD	0.0675	0.1632	0.1679	0.1681	0.1701	0.1701	0.1684	0.0000	0.0029	0.1674
UC	0.0677	0.1632	0.1681	0.1681	0.1701	0.1701	0.1684	0.0000	0.0029	0.1674
conSSD	0.0000	0.0003	0.1548	0.0000	0.0054	0.0041	0.0165	0.0000	0.0013	0.1146
conCSSD	0.0675	0.0043	0.0080	0.0005	0.0059	0.0057	0.0016	0.0000	0.0029	0.0040
PC	0.0677	0.0548	0.0139	0.1473	0.0076	0.0076	0.1682	0.0000	0.0029	0.0070

Table 4. Shift in the steady-state mass of undesirable states of the mutated mammalian cell-cycle network for different control genes and different control policies

Control	E2F	CycE	CycA	Cdc20	Cdh1	UbcH10	CycB
SSD	0.1972	0.0573	0.0334	0.0565	0.0835	0.0122	0.0598
CSSD	0.1972	0.0573	0.0335	0.0836	0.0837	0.0138	0.0843
UC	0.1972	0.0573	0.0335	0.0836	0.0837	0.0138	0.0843
conSSD	0.0374	0.0207	0.0329	0.0119	0.0327	0.0001	0.0244
conCSSD	0.0571	0.0573	0.0332	0.0332	0.0506	0.0118	0.0254
PC	0.1857	0.0573	0.0332	0.0565	0.0509	0.0120	0.0481

willing to take. The seemingly intuitive choice of directly perturbing WNT5A to control WNT5A (first column) only produces the best result with the conSSD policy. This is because of feedback in the network. Not only can a direct approach absent constraint produce unwanted phenotypes, a direct approach may not even achieve the desired reduction in long-run probability mass of the target gene on account of complicated feedback loops.

The aggregated mass of undesirable states for the uncontrolled mammalian cell cycle network is $\sum_{j \in U} \pi_j = 0.2012$. Table 4 gives the amount of shift made by all control policies for different control gene candidates. The table shows that choosing E2F as the control gene induces the maximum shift under different control policies. Thus, E2F can be considered as a potential control gene for intervention. It is also evident that the difference between the performances of PC optimal policies and conCSSD is significant.

4 CONCLUDING REMARKS

Heretofore, two external control approaches have been taken to shift the steady-state mass of a GRN: (i) use a subjectively defined cost function for which desirable shift of the steady-state mass is a by-product and (ii) use heuristics to design a greedy algorithm. The present article uses a linear programming approach to optimally shift the steady-state mass and therefore outperforms both of the preceding approaches. Moreover, it does with less computational overhead and therefore can be applied to larger networks. Owing to its generality, the same basic LP structure is used for both unconstrained and constrained optimization problems. Thus, at least in the case of

Markovian GRNs, in particular, PBNs, previously proposed methods can be abandoned whenever the goal is to optimally shift the steady-state mass.

Conflict of Interest: none declared.

REFERENCES

- Altman, E. (1999) *Constrained Markov Decision Processes*. Chapman & Hall/CRC, Boca Raton, FL, USA.
- Bittner, M. *et al.* (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature*, **406**, 536–540.
- Brun, M. *et al.* (2005) Steady-state probabilities for attractors in probabilistic Boolean networks. *Signal. Process.*, **85**, 1993–2013.
- Ching, W.K. *et al.* (2009) Optimal control policy for probabilistic boolean networks with hard constraints. *IET Syst. Biol.*, **3**, 90–99.
- Cho, G.E. and Meyer, C.D. (2001) Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra Appl.*, **335**, 137–150.
- Datta, A. *et al.* (2003) External control in Markovian genetic regulatory networks. *Mach. Learn.*, **52**, 169–191.
- Datta, A. and Dougherty, E.R. (2006) *Introduction to Genomic Signal Processing with Control*. CRC Press.
- Derman, C. (1970) *Finite State Markovian Decision Processes*. Academic Press, New York.
- Dissanayake, S.K. *et al.* (2008) Wnt5a regulates expression of tumor-associated antigens in melanoma via changes in signal transducers and activators of transcription 3 phosphorylation. *Cancer Res.*, **68**, 10205–10214.
- Dougherty, E.R. and Datta, A. (2005) Genomic signal processing: diagnosis and therapy. *IEEE Signal Process. Mag.*, **22**, 107–112.
- Dougherty, E.R. *et al.* (2010) Stationary and structural control in gene regulatory networks: basic concepts. *Int. J. Syst. Sci.*, **41**, 5–16.
- Faryabi, B. *et al.* (2007) On approximate stochastic control in genetic regulatory networks. *IET Syst. Biol.*, **1**, 361–368.
- Faryabi, B. *et al.* (2009) Intervention in context-sensitive probabilistic Boolean networks revisited. *EURASIP J. Bioinform. Syst. Biol.*, **2009**, Article ID 360864.

- Faure, A. *et al.* (2006) Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, **22**, e124–e131.
- Ghaffari, N. *et al.* (2010) A CoD-based reduction algorithm for designing stationary control policies on Boolean networks. *Bioinformatics*, **26**, 1556–1563.
- Ivanov, I. *et al.* (2007) Dynamics preserving size reduction mappings for probabilistic Boolean networks. *IEEE Trans. Signal Process.*, **55**, 2310–2322.
- Ivanov, I. *et al.* (2010) Selection policy-induced reduction mappings for Boolean networks. *IEEE Trans. Signal Process.*, **58**, 4871–4882.
- Kallenberg, L.C.M. (1983) *Linear Programming and Finite Markovian Control Problems*. Mathematisch Centrum, Amsterdam.
- Kauffman, S.A. (1993) *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, NY, USA.
- Law, A. and Wong, C. (2010a) Stanniocalcin-2 is a Hif-1 target gene that promotes cell proliferation in hypoxia. *Exp. Cell Res.*, **316**, 466–476.
- Law, A. and Wong, C. (2010b) Stanniocalcin-2 promotes epithelial-mesenchymal transition and invasiveness in hypoxic human ovarian cancer cells. *Exp. Cell Res.*, **316**, 3425–3434.
- Megiddo, N. (1987) On the complexity of linear programming. In: *Advances in Economic Theory: Fifth World Congress*. Cambridge University Press, New York, NY, USA, pp. 225–268.
- Pal, R. *et al.* (2006) Optimal infinite-horizon control for probabilistic Boolean networks. *IEEE Trans. Signal Process.*, **54**, 2375–2387.
- Pal, R. *et al.* (2008) Robust intervention in probabilistic Boolean networks. *IEEE Trans. Signal Process.*, **56**, 1280–1294.
- Pal, R. *et al.* (2009) Bayesian robustness in the control of gene regulatory networks. *IEEE Trans. Signal Process.*, **57**, 3667–3678.
- Qian, X. and Dougherty, E.R. (2008) Effect of function perturbation on the steady-state distribution of genetic regulatory networks: optimal structural intervention. *IEEE Trans. Signal Process.*, **56**, 4966–4976.
- Qian, X. and Dougherty, E.R. (2012) Intervention in gene regulatory networks via phenotypically constrained control policies based on long-run behavior. *IEEE ACM Trans. Comput. Biol.*, **9**, 123–136.
- Qian, X. *et al.* (2009) Intervention in gene regulatory networks via greedy control policies based on long-run behavior. *BMC Syst. Biol.*, **3**, 61.
- Qian, X. *et al.* (2010) State reduction for network intervention with probabilistic Boolean networks. *Bioinformatics*, **26**, 3098–3194.
- Ross, K.W. (1989) Randomized and past-dependent policies for Markov decision processes with multiple constraints. *Oper. Res.*, **37**, 474–477.
- Seneta, E. (1988) Perturbation of the stationary distribution measured by ergodicity coefficients. *Adv. Appl. Probab.*, **20**, 228–230.
- Shmulevich, I. *et al.* (2002a) Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, **18**, 261–274.
- Shmulevich, I. *et al.* (2002b) Gene perturbation and intervention in probabilistic Boolean networks. *Bioinformatics*, **18**, 1319–1331.
- Shmulevich, I. *et al.* (2002c) Control of stationary behavior in probabilistic Boolean networks by means of structural intervention. *J. Biol. Syst.*, **10**, 431–445.
- Vahedi, G. *et al.* (2008) Intervention in gene regulatory networks via a stationary mean-first-passage-time control policy. *IEEE Trans. Biomed. Eng.*, **55**, 2319–2331.
- Weeraratna, A.T. *et al.* (2002) Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell*, **1**, 279–288.
- Xiao, Y. and Dougherty, E.R. (2007) The impact of function perturbations in Boolean networks. *Bioinformatics*, **23**, 1265–1273.
- Yousefi, M.R. *et al.* (2012) Optimal intervention strategies for therapeutic methods with fixed-length duration of drug effectiveness. *IEEE Trans. Signal Process.*, **60**, 4930–4944.