SYSTEMS MEDICINE:

AN INTEGRATED APPROACH WITH DECISION MAKING PERSPECTIVE

A Dissertation

by

BABAK FARYABI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2009

Major Subject: Electrical Engineering

SYSTEMS MEDICINE:

AN INTEGRATED APPROACH WITH DECISION MAKING PERSPECTIVE

A Dissertation

by

BABAK FARYABI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Edward R. Dougherty
	Aniruddha Datta
Committee Members,	Jean-Francois Chamberland
	N. Sivakumar
Head of Department,	Costas N. Georghiades

August 2009

Major Subject: Electrical Engineering

ABSTRACT

Systems Medicine:

An Integrated Approach with Decision Making Perspective. (August 2009)

Babak Faryabi, B.S., Sharif University of Technology;

M.S., Sharif University of Technology

Co–Chairs of Advisory Committee: Dr. Edward R. Dougherty
Dr. Aniruddha Datta

Two models are proposed to describe interactions among genes, transcription factors, and signaling cascades involved in regulating a cellular sub-system. These models fall within the class of Markovian regulatory networks, and can accommodate for different biological time scales. These regulatory networks are used to study pathological cellular dynamics and discover treatments that beneficially alter those dynamics. The salient translational goal is to design effective therapeutic actions that desirably modify a pathological cellular behavior via external treatments that vary the expressions of targeted genes. The objective of therapeutic actions is to reduce the likelihood of the pathological phenotypes related to a disease. The task of finding effective treatments is formulated as sequential decision making processes that discriminate the gene-expression profiles with high pathological competence versus those with low pathological competence. Thereby, the proposed computational frameworks provide tools that facilitate the discovery of effective drug targets and the design of potent therapeutic actions on them. Each of the proposed system-based therapeutic methods in this dissertation is motivated by practical and analytical considerations. First, it is determined how asynchronous regulatory models can be used as a tool to search for effective therapeutic interventions. Then, a constrained intervention

method is introduced to incorporate the side-effects of treatments while searching for a sequence of potent therapeutic actions. Lastly, to bypass the impediment of model inference and to mitigate the numerical challenges of exhaustive search algorithms, a heuristic method is proposed for designing system-based therapies. The presentation of the key ideas in method is facilitated with the help of several case studies.

To My Love, Golnaz.

ACKNOWLEDGMENTS

I would like to especially thank my advisor and co-chair, Dr. Edward Dougherty, from whom I learned a lot about life in general and research in particular. His guidance and great vision were instrumental in the completion of this dissertation.

I am indebted to my other co-chair, Dr. Aniruddha Datta, for his invaluable encouragement and support throughout my research.

I am also deeply grateful to Dr. Jean-Francios Chamberland for his guidance and informative discussions. He always gave a patient hearing to my ideas and encouraged me to enhance my understanding through learning fundamental concepts.

I am thankful to my parents for their immense support, unconditional love and inspiration throughout my life. Last and most of all, my debt to my wife is beyond words. She was the most reliable source of emotional solace and intellectual support throughout my graduate study.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

In biology, there are numerous examples where the (in)activation of one gene or protein can lead to a certain cellular functional state or phenotype. For instance, in a stable cancer cell line, the reproductive cell cycle is repeated and cancerous cells proliferate with time in the absence of intervention. One can use the $p53$ gene if the intervention goal is to push the cells into apoptosis, or programmed cell death, to arrest the cell cycle. The $p53$ gene is one of the most well-known tumor suppressor gene, encoding a protein that regulates the expression of several genes such as *Bax* and *Fas/Apo*1, whose function is to promote apoptosis [1] [2]. In cultured cells, extensive experimental results indicate that when $p53$ is activated, e.g. in response to radiation, it leads to cell growth inhibition or cell death [3]. The $p53$ gene is also used in gene therapy, where the target gene ($p53$ in this case) is cloned into a viral vector. The modified virus serves as a vehicle to transport the $p53$ gene into tumor cells to generate intervention [4, 5]. As this and many other examples suggest, it is prudent to use gene regulatory models to design therapeutic interventions that expediently modify the cell's dynamics via external signals. These system-based intervention methods can be useful in identifying potential drug targets and discovering treatments to disrupt or mitigate the aberrant gene functions contributing to the pathology of a disease.

In [6] and [7], few methods to design therapeutic interventions are discussed. Some of these methods are intended to reduce the likelihood of the gene-expression profiles associated with aberrant cellular functions via manipulation of a control gene. In a nutshell, whenever changing the expression level of a control gene is perceived

This dissertation follows the style of *IEEE Journal of Selected Topics in Signal Processing.*

as a therapeutic option, e.g. $p53$ in the above scenario, these system-based therapies identify for the most effective sequence of such changes to beneficially alter cell dynamics. The resulting intervention strategy specifies the appropriate expression of the control gene in order to reduce the likelihood of pathological cellular functions.

In the case of a cancerous tumor, the objective of treatment could be to diminish the long-run likelihood of metastasis. In this scenario, one may consider the correlation between metastasis and the abundances of messenger RNA for certain genes. For instance, the abundance of messenger RNA for the gene $Wnt5a$ has been found to be highly discriminating between cells with properties typically associated with high versus low metastatic competence [8]. The messenger RNA level of the gene $Wnt5a$ can therefore be used to annotate profiles of gene expressions as desirable and undesirable. One partition of gene-expression profiles corresponds to high, while the other to low, metastatic competence. Having defined a cost function to discriminate between the two sets of gene-expression profiles, the task of finding an effective intervention strategy can be mathematically formulated as a sequential decision making problem for a pre-defined cost of intervention. The objective of the decision maker is to identify a strategy that minimizes a well-defined function of the accumulated cost over time. Such a strategy can be seen as a system-based treatment to avoid undesirable gene-expression profiles contributing to the pathology of a disease.

These intervention design methods utilize the information in an inferred regulatory network to devise a sequence of actions on control genes. Thus, modeling the regulatory processes in the cells is a first step toward a systematic approach toward treatment discovery. Constructing complex regulatory models that finely describe the interactions among biological components related to a disease requires precise understanding of the underlying biological processes and extensive amounts of experimental data. Although genetically pathological cells, such as cancer cells, have been

extensively studied, we still lack sufficient knowledge and data sets to construct such complex models. Meanwhile, it is equally important, especially from a translational perspective, to discover effective drug targets and devise therapeutic strategies with the help of simpler regulatory models, such as Markovian regulatory models.

In Chapter II, we describe two classes of Markovian networks: synchronous and asynchronous. Major efforts have been focused on devising intervention strategies that affect the dynamics of synchronous Markovian networks. The effect of an intervention strategy that is beneficial in the short-term may wear out over time. Thus, it is important to look for intervention strategies that consider the long-run effects. In the framework of synchronous networks, the theory of infinite-horizon Markovian decision processes can be employed to find optimal intervention strategies with respect to the defined objective functions [6]. An optimal strategy determines the actions to be taken using the external signal in response to each gene-expression profile.

Formulating the problem of intervention in a regulatory network as a classical infinite-horizon decision making process is introduced in Chapter IV. We refer to this intervention method as *classical intervention* throughout this dissertation. A classical intervention scheme introduces an elegant analytical framework that may be instrumental to enhance our understanding of treatment discovery.

Despite its conceptual benefits, classical intervention fails to address many practical and technical issues. In this volume, we extend the classical framework in several directions to improve system-based intervention schemes. To achieve this goal, we have envisaged a number of objectives for which both methodology and techniques must be improved. The proposed analytical methods provide insight into the design of effective therapeutic interventions, and strive to address some of the practical concerns that are brought up by medical practitioners. In each chapter, we strive to highlight the objectives behind each proposed method and explain it via a biological

case study. Before proceeding further, we briefly discuss the motivation behind each of the schemes discussed in this volume.

The scheme to update the gene values in a regulatory model plays a crucial role in its ability to describe the dynamics of gene interactions and thereby influences the effectiveness of designed therapies. In Chapter II, we describe context-sensitive probabilistic Boolean networks. These are a class of synchronous Markovian regulatory networks that allow the incorporation of uncertainty into the inter-gene relationships. Synchronous Boolean networks are a class of discrete-time discrete-space Markovian regulatory networks in which all the elements in the model are assumed to be updated simultaneously [9–11]. In the same chapter, we derive a closed-form representation for the transition probability matrix that describes the dynamics of this category of regulatory models.

From a biological perspective, interactions among genes and proteins causing various processes, such as transcription, translation, and degradation, occur over a wide range of time-scales [12, 13]. In a synchronous model, the tacit assumption is that asynchronous updating will not unduly alter the presented biological properties central to the application of interest. This assumption may not generally hold. Various potential issues with synchronous networks have been raised [14–16].

These observations motivate us to study intervention based on discrete state-space models that can capture timing information in gene interactions. An asynchronous Markovian regulatory model, suited to our intervention objective, should poss four characteristics: (1) it should be inferable from the empirical time-course measurements; (2) it should accurately represent relations among macromolecules of interest; (3) it should enable us to analytically study the temporal behavior of relevant phenomena; and (4) the model should be appropriate for the study of therapeutic intervention. With these conditions in mind, we propose two asynchronous

Markovian regulatory networks in Chapter II. In the first asynchronous framework, the updating period of each gene is fixed, but can differ from one gene to another. The second asynchronous regulatory model introduces asynchronism relative to the state-space of gene-expression profiles. This approach is more suitable from an inference perspective.

After introducing these models in Chapter II, we propose computational tools that search for effective therapeutic interventions using the information in these models. In Chapter V, we show how the decision making problem in an asynchronous regulatory network can be reformulated as a synchronous decision making problem similar to the one formulated in Chapter IV. Hence, it would be sufficient to study the intervention techniques described in the later chapters only in the framework of synchronous models. These asynchronous models can potentially provide more effective intervention strategies, depending on our ability to perform satisfactory inference.

Sequential decision making techniques can be categorized into two classes. The first class of schemes, such as classical intervention, requires exact optimization of a cost function by the decision maker to find an effective treatment within the space of all possible treatments. It is well-known that the exact solution of such a search problem is not robust relative to inaccuracy of the underlying network. The procedures for regulatory network inference are prone to modeling errors. They suffer from insufficient empirical measurements and computational complexity [17]. In addition, the computational complexity of the optimization at the heart of the classical intervention prohibits its use with regulatory models possessing large numbers of components. To bypass the impediment of model inference and to mitigate the numerical problems associated with the exact optimization approach, heuristic schemes can be used to design system-based therapies.

Some heuristic intervention methods estimate insightful statistics of the regu-

latory network from empirical measurements and utilizes these statistics to greedily search the space of all strategies for an effective one. Typically, effective heuristic methods provide lower computational complexity, and also adapt to changes in the underlying biological system.

In Chapter VII, we introduce a heuristic intervention scheme based on reinforcement learning. Given the cost structure, the reinforcement intervention estimates the statistics of the cost function and uses this information to learn an effective strategy. This heuristic strategy progressively improves its performance as more empirical measurements become available.

Besides confronting inferential, complexity, and robustness problems, which are essentially engineering issues, one also needs to take into account practical medical issues. Consider the fact that medicine is able to exploit the biochemical differences between bacteria and human cells so as to achieve toxic drug concentrations in the former while sparing the latter. This selectivity largely contributes to the success in treating bacterial infections. Unfortunately, such high selectivity is at present elusive in the treatment of human cancers. Hence, great efforts are required to determine dose schedules that maximize the benefit to toxicity ratio in cancer therapy [18].

Dose intensity is a measure of treatment delivery that looks at the amount of drug delivered per unit of time. To mitigate the detrimental side effects of a treatment in general, we should account for dose intensity in a system-based intervention method. A therapeutic intervention should avoid undesirable gene-expression profiles while accounting for the quantity and/or frequency of applied drugs. A higher drug dose intensity can be delivered by increasing the dose per cycle (dose escalation) or by reducing the interval between cycles (dose density).

Dose intensity can be regulated by the number of interventions in a therapeutic strategy. A treatment based on estrogen is often used by women after menopause

to alter their accelerated aging trend. The amount of estrogen received during this treatment should not exceed a threshold. An overdose may increase the chance of developing breast and ovarian cancers. While this phenomenon is not fully understood, it is conceivable that estrogen therapy may have side effects on gene regulation. Estrogen generates two types of complexes through binding to two classes of receptors. The generated complexes are transported into the nucleus to bind to the enhancer elements on the target genes with the help of a coactivator. The coactivator is also required for efficient transcriptional regulation by estrogen. This function, in cooperation with a coactivator acts like a transcription factor, affecting target genes such as the *PENK* gene [19]. Two types of receptors are competing for binding to the estrogen received via treatment [20]. The first type of complex binds DNA better but performs less efficiently to bind the coactivator; the second type of complex binds the coactivator better but performs poorly when binding DNA. When the level of estrogen is below a threshold, there is no competition for DNA binding. Hence, the second type of complex also binds DNA and activates the downstream target gene *PENK*, with the help of the coactivator. However, when the estrogen level is high, both types of complexes exist at high concentrations and the second type of complex binds and depletes the coactivator. The level of coactivator available to complex type one drops. Hence, the complex type one does not have necessary coactivator, and has a small chance to bind to DNA and causes activation of gene *PENK*. If the *PENK* gene plays a role in tumor suppression, for instance, then this could explain why high levels of estrogen have a tumorigenic effect. An appropriate treatment strategy mitigates this problem by bounding the expected number of treatments received by a patient and, as a result, limits the dose intensity of estrogen.

In general, the likelihood of eradicating pathological cell functions is maximized by delivering the most effective dose intensity of a drug whose toxicity can be tol-

erated by the patient. In addition, levels higher than a certain concentration may not increase the killing rate of cancer cells. The dose intensity of a drug is directly related to the number of interventions in a therapeutic strategy. *Constrained intervention* is introduced in Chapter VI to incorporate the aforementioned concern in the system-based therapy design paradigm. Using constrained intervention methods, we seek an effective regulatory treatment that reduces the likelihood of visiting undesirable gene-expression profiles in the long run while providing an upper bound on the expected number of interventions a patient may receive [21].

CHAPTER II

MARKOVIAN REGULATORY NETWORKS *

Systems biology studies the multivariate interaction among biological components, e.g. genes and proteins. From a translational perspective, the ultimate objective of regulatory modeling is to use the model to design intervention strategies that beneficially alter the dynamics of inter-gene interactions, for instance, to reduce the likelihood of states favorable to metastasis in cancer cells.

We study regulatory intervention in the framework of *Markovian regulatory networks*. Constructing complex regulatory models that finely describe the interactions among biological components related to a disease requires precise understanding of the underlying biological processes and extensive amounts of experimental data. Although genetically pathological cells, such as cancer cells, have been extensively studied, we still lack sufficient knowledge and data sets to construct such complex models. Meanwhile, it is equally important, especially from a translational perspective, to discover effective drug targets and devise therapeutic strategies with the help of simpler regulatory models, such as Markovian regulatory models. To this end, *probabilistic Boolean networks* [22], which compose a class of discrete-time discrete-space Markovian regulatory networks have been utilized to devise system-based therapeutic interventions [6, 7]. These classes of rule-based models, which allow the incorpora-

tion of uncertainty into inter-gene relationships, are probabilistic generalizations of classical *Boolean networks* [9–11].

In rule-based regulatory networks, such as Boolean networks, a *regulatory graph* defines the multivariate interactions among the components. From here on, we use the term *gene* in place of any general biological components, e.g. genes and transcription factors, involved in a regulatory network. The vertices of a regulatory graph are the *genes*. A directed edge starts from a predictor vertex and ends at an influenced vertex. All the vertices directly connected to a gene are its predictors. A regulatory rule defines the multivariate effects of predictors on the vertex. The gene values are selected from a set of possible quantization levels to facilitate the modeling of gene interactions by logical rules. The discrete formalism of rule-based regulatory networks is plausible for many classes of biological systems. Strong evidences suggest that the input-output relations of regulatory interactions are sigmoidal and can be well approximated by step functions [23, 24].

If gene values are quantized to two levels, then the rule-based networks are described by a collection of Boolean functions, with 0 or 1 meaning genes are OFF or ON, respectively. Ternary quantization arises when we consider individual genes to be down-regulated, up-regulated, or invariant. This situation commonly occurs with cDNA microarrays, where a ratio is taken between the expression values on the test channel (red) and the base channel (green) [25]. In this thesis, we develop the methodology for $d = 2$, so that gene values are either 0 or 1. The methodology can be extended to other finite quantization levels, albeit at the expense of tedious mathematical expressions. All the binary operations in this dissertation would need to be replaced by case statements and the perturbation process should be articulated on a case by case basis.

Figure 1 shows the regulatory graph of a hypothetical three-gene network. There

is a unidirectional relation between genes x_1 and x_2. The relation between genes x_2 and x_3 is bidirectional.

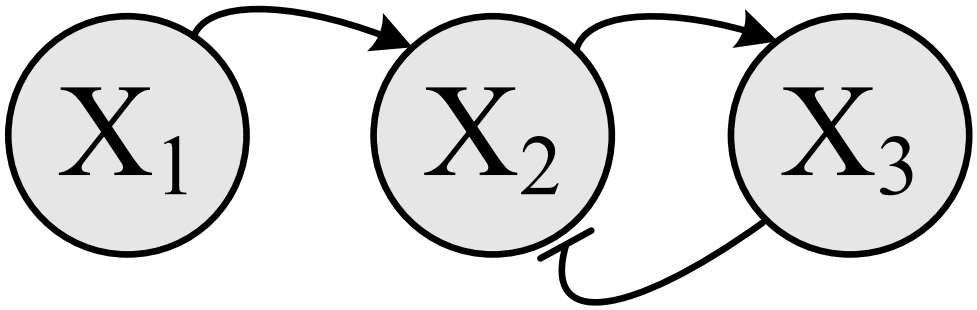


Fig. 1. Presentation of a regulatory graph for an arbitrary 3-gene Boolean network.

To completely specify dynamics of inter-gene interactions, we need to adopt an updating scheme in a class of regulatory networks. The choice of the updating scheme plays a crucial rule in the dynamical behavior of the network. Given the updating scheme, we can depict the dynamical evolution of genes by translating the information of the regulatory graph and the regulatory rules into an *oriented graph*. The vertex of an oriented graph represents a *state*, which is the vectored values of all the genes at a given time. An edge traverses from one state to another if a transition can occur in the direction of the edge from one vertex to the other.

The choice of the updating scheme plays a crucial role in the dynamical behavior of the network. In Boolean networks, the values of genes are updated synchronously at equally distant updating epochs. For instance, Figure 2 shows the oriented graph corresponding to the regulatory graph in Figure 1. According to this oriented graph, whenever the aggregate value of the three genes in the network is ($x_1 = 0$, $x_2 = 1$, $x_3 = 0$) and if all the genes are updated synchronously, then the next state is ($x_1 = 0$, $x_2 = 0$, $x_3 = 1$).

A regulatory graph is a static representation of interactions among biological

components, whereas an oriented graph shows the dynamics of interactions among these components. A key point concerning Boolean networks is that, in the long run, the network will settle into an *attractor cycle* (e.g "000", or "110" and "101" in Figure 2), meaning that the network will endlessly cycle through some set of states. We can practically observe timing information related to the dynamical representation of biological component interactions, that is, timing relative to the oriented graph.

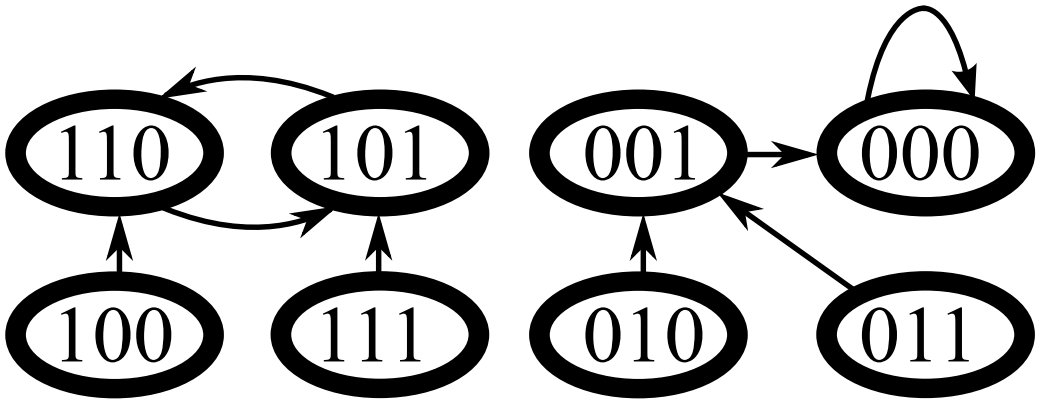


Fig. 2. Presentation of the oriented graph corresponding to a regulatory graph for the 3-gene Boolean network in Figure 1.

To incorporate the effects of latent variables outside the model, whose behaviors influence regulation within the system, stochasticity is introduced into the Boolean model by allowing several possible regulatory functions for each gene and allowing random modification of the genes [22]. The resulting model is called a *probabilistic Boolean network* (PBN), where the terminology *Boolean* refers to the logical character of the relations, not that they are necessarily binary. If the regulatory functions are allowed to change at every time point, then the PBN is said to be *instantaneously random* [22]. On the other hand, in a *context-sensitive* PBN, function updating only occurs at time points selected by a Bernoulli random switching process [26, 27]. In essence, a PBN is composed of a collection of networks (oriented graphs); between

switches it acts like one of the constituent networks (oriented graphs), each being referred to as a *context*. The switching frequency of the context differentiates the instantaneously random PBN from the context-sensitive PBN. The PBN model also allows random perturbation of genes at each updating instant. By definition, the attractors of a PBN consist of the attractors of its constituent contexts.

Under appropriate assumptions, a Markov chain models the dynamical behavior of an instantaneously random PBN [28]. Hence, the oriented graph of an instantaneously random PBN can be represented by a Markov chain. The associated Markov chain to an instantaneously random PBN has been used to develop effective intervention strategies. The transition probability distributions associated with a PBN act on its states and describe their trajectories over time.

The transition probability distributions of an instantaneously random PBN are derived in [28]. In this chapter, we derive a closed-form representation for transition probability distributions among the states of a context-sensitive PBN. This expression of transition probability distribution is in concert with the original definition of this class of Markovian networks in [27] and [26].

From a biological perspective, interactions among genes causing transcription, translation, and degradation occur over a wide range of time-scales. Earlier studies suggest that asynchronously updating the genes alters the global behavior of synchronous networks due to the change in their oriented graph, which models the dynamics of the system [14–16]. Synchronous abstraction is used under the implicit assumption that asynchronous updating will not unduly alter the properties of a system central to the application of interest [14]. Clearly, some properties will be altered. For instance, in Figure 2, if all the genes are not simultaneously updated, then state "001" may transition to some other state instead of transitioning to the attractor state "000".

In this chapter, we further relax the synchronism assumption in Markovian regulatory networks such as PBN and consider asynchronous networks. This involves defining asynchronous Markovian regulatory networks and treating the intervention problem in the framework of asynchronous processes as covered in Chapter V.

Two new asynchronous Markovian networks and methods to derive effective intervention strategies for each one of them are proposed. The first model introduces asynchronism into Markovian regulatory networks by updating each gene based on its timing information period. This approach of introducing asynchronism into networks extends the previously favored approach of studying asynchronism in regulatory networks. The proposed model, called a *deterministic-asynchronous probabilistic Boolean network*, is an extension of context-sensitive PBNs in which the time scales of various biological updates can be different.

The second model extends synchronous networks by considering asynchronism at the state level. Assuming asynchronism at the gene level for Boolean networks has practical and theoretical impediments that may prevent independent gene updating to serve as a basis for designing effective therapeutic intervention strategies [14, 29, 30]. To this end, we propose *semi-Markov asynchronous regulatory networks*. In this regulatory model the asynchronism is at the gene-expression profile level.

In *semi-Markov asynchronous regulatory networks*, the empirically measurable timing information of biological systems is incorporated into the model. This temporal information determines the typical time delay between transitions from one gene-expression profile to another. Since the order of updating genes and their relative time delays depend on the levels of other regulatory components, time-course data enable the estimation of inter-transition times between gene-expression profiles, not the updating time of each gene. It is then natural to introduce asynchronism at the gene-expression profile level.

This approach is resourceful from a translational perspective. While the physical evolution of the biological network occurs over continuous time, the PBN records only the state transitions and contains no information on the time between the individual transitions. The PBN model inherits this property from the Boolean model, which it generalizes. Hence, the problem can be explained in the framework of the Boolean model. Figure 3 shows two continuous-time realizations that are equivalent from the point of view of the model of Figure 2. In both figures 3(a) and 3(b), the initial state is "100". We observe the evolution "100"→"010"→"001", at which point there are no other changes because "001" is an attractor of the network. While equivalent from the perspective of the Boolean model, from the perspective of continuous time observation, the realizations of Figures 3(a) and 3(b) are not the same. For instance, in the second realization, the sojourn time in state "010" is much longer than in the first realization. This may be of no concern if we are only interested in tracking the state transitions. On the other hand, suppose we are considering intervention and penalizing undesirable phenotypes. Then, if "010" is an undesirable state, the penalty should be greater in the second realization; that is, the penalty needs to account for the sojourn time in a state. This problem has been addressed in the framework of asynchronous Markovian regulatory networks by considering the process to be defined over continuous time.

A. Synchronous Markovian Regulatory Networks

In this Section, we review the definition of a context-sensitive PBN [27]. We derive a new expression for the transition probability matrix that specifies its oriented graph. This expression for the transition probability matrix is in concert with the original definition of context-sensitive PBNs [27].

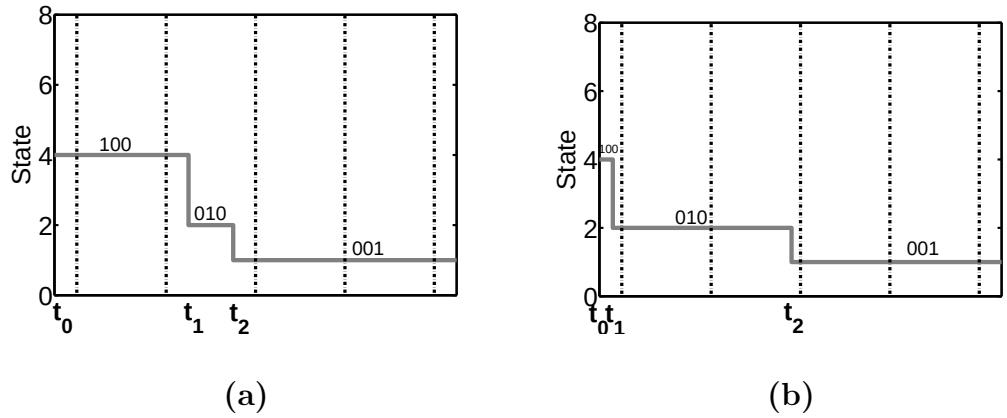


Fig. 3. Two realizations of trajectories for the oriented graph in Figure 2.

PBNs are a class of discrete-time discrete-space Markovian regulatory networks that allow the incorporation of uncertainty into the inter-gene relationships. In a PBN model, gene values are quantized into some finite range. The values are updated synchronously at each time step according to regulatory functions. Stochastic properties are introduced into the model by allowing several possible regulatory functions for each gene and allowing random modification of the activity factors. In essence, a PBN is composed of a collection of networks; between switches it acts like one of the constituent networks, each being referred to as a *context*. The switching frequency of the context differentiates the instantaneously random PBN [22] from the context-sensitive PBN [26, 27].

The context switching that occurs at every time step in an instantaneously random PBN corresponds to changing the wiring diagram of the system at every instant. In contrast, context-sensitive PBNs can better represent the stability of biological systems by capturing the period of sojourning in constituent networks [26]. Hence, this class of models is more suitable for the analysis of gene regulation and the design of intervention methods. To formulate the problem of intervention in a context-sensitive

PBN, a transition probability matrix must be derived. This transition matrix acts on the possible states of the system. Once this is accomplished, the task of finding the most effective intervention strategy can be formulated as a classical sequential decision making problem. The latter is covered in Chapter IV, here we compute the transition probability matrix of a context-sensitive PBN.

It is evident that the intervention strategy specified by a sequential decision maker is directly affected by the form of the transition probability matrix associated with a PBN. For an instantaneously random PBN, the state consists of a gene-expression profile; while for a context-sensitive PBN, the state includes a gene-expression profile and a context. The effectiveness of an intervention strategy depends, partly, on how accurate the model represents the reality of the underlying pathological cellular functions. It is therefore important to adopt a model that captures the subtleties of the biological system of interest. In the framework of context-sensitive PBNs, this entails defining a transition probability matrix that is an accurate representation of system dynamics [31].

1. Context-Sensitive Probabilistic Boolean Networks: Definition

A context-sensitive probabilistic Boolean network consists of a sequence $V = \{x_i\}_{i=1}^{n}$ of n nodes, where $x_i \in \{0, 1\}$, and a sequence $\{\boldsymbol{f}_l\}_{l=1}^{k}$ of vector-valued functions defining constituent networks. In the framework of gene regulation, each element x_i represents the expression value of a gene. It is common to abuse terminology by referring to x_i as the ith gene. Each vector-valued function $\boldsymbol{f}_l = (f_{l1}, \ldots, f_{ln})$ represents a constituent network of the context-sensitive PBN. The function $f_{li} : \{0, 1\}^n \to \{0, 1\}$ is the predictor of gene i, whenever context l is selected.

At each updating epoch a random variable determines whether the constituent network is switched or not. The switching probability q is a system parameter. If

the context remains unchanged, then the context-sensitive PBN behaves like a fixed Boolean network where the values of all the genes are updated synchronously according to the current constituent network. On the other hand, if a switch occurs, then a constituent network is randomly selected from $\{\boldsymbol{f}_l\}_{l=1}^{k}$ according to the selection probability distribution $\{r_l\}_{l=1}^{k}$. Once the predictor function $\boldsymbol{f}_l$ is determined, the values of the genes are updated using the new constituent network; that is, according to the rules defined by $\boldsymbol{f}_l$.

We focus on context-sensitive PBNs with perturbations, meaning that each gene may change its value with small probability p at each epoch. If $\gamma_i(t)$ is a Bernoulli random variable with parameter p and the random vector $\boldsymbol{\gamma}$ at instant t is defined as $\boldsymbol{\gamma}(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t))$, then the value of gene i is determined at each epoch t by

$$\begin{aligned} x_i(t+1) =& \mathbf{1}(\boldsymbol{\gamma}(t+1) \neq \mathbf{0})\,(x_i(t) \oplus \gamma_i(t+1)) \\ &+ \mathbf{1}(\boldsymbol{\gamma}(t+1) = \mathbf{0}) f_{li}(x_1(t), \ldots, x_n(t)), \end{aligned} \tag{2.1}$$

where operator $\oplus$ is component-wise addition in modulo two and f_{li} is the predictor of gene i according to the current context of the network l. Such a perturbation captures the realistic situation where the activity of a gene is subject to random alteration. As we will see in the next section, in addition, it guarantees that the Markov chain associated to the oriented graph of a PBN attain a unique steady-state distribution.

The gene-activity profile (or GAP) is an n-digit binary vector $\boldsymbol{x}(t) = (x_1(t), \ldots, x_n(t))$, giving the expression values of the genes at time t. Here, $x_i(t) \in \{0, 1\}$. We denote the set of all possible GAPs by $\mathcal{X} = \{0, 1\}^n$.

The dynamic behavior of a context-sensitive PBN can be modeled by a Markov chain whose states are ordered pairs consisting of a constituent network κ and a GAP $\boldsymbol{x}$. The evolution of the context-sensitive PBN can therefore be represented using a

stationary discrete-time equation

$$\boldsymbol{z}(t+1) = f(\boldsymbol{z}(t), w(t)) \qquad t = 0, 1, \ldots \tag{2.2}$$

where state $\boldsymbol{z}(t)$ is an element of the state-space $\mathcal{Z} = \{(\kappa, \boldsymbol{x}) : \kappa \in \{1, \ldots, k\}, \boldsymbol{x} \in \mathcal{X}\}$. The disturbance $w(t)$ is the manifestation of uncertainties in the biological system, due either to context switching or a change in the GAP resulting from a random gene perturbation.

It is assumed that both the gene perturbation distribution and the network switching distribution are independent and identically distributed over time.

We note that a bijection can be drawn between the gene-activity profile $\boldsymbol{x}(t)$ or the states $\boldsymbol{z}(t)$ and their decimal representations $x(t)$ and $z(t)$ based on their binary expansion. The integers $x(t)$ and $z(t)$ take values in $X = \{0, 1, \ldots, 2^n - 1\}$ and $Z = \{0, 1, \ldots, k \times 2^n - 1\}$, respectively. These decimal representations facilitate the depiction of our numerical results in the next chapters.

2. Context-Sensitive Probabilistic Boolean Networks: Transition Probability Matrix

The oriented graph of a context-sensitive PBN can be represented by a Markov chain [31]. The transition probability distributions associated with a context-sensitive PBN act on its states and describe their trajectories over time.

The latter can be found by observing that two mutually exclusive events may occur at any epoch: the current context of the network remains the same for two consecutive instants, or the context of the network changes to a new one at the time instant $t+1$. Moreover, the context may remain unchanged in two mutually exclusive ways: the binary switching variable is 0, which means that no change is possible, or the binary switching variable is 1 and the current network is picked again through

random selection [26, 27]. In particular, when the switching variable is 1, a new context is selected independent of the current system state. Thus, the same network can be active twice in a row. This interpretation of switching the context in a PBN is in concert with the original definition of context-sensitive PBNs in [27]. Before proceeding, we note that transitioning was defined differently in [32] and [33], where it was assumed that, whenever the switching variable is 1, a change of context must occur; the result being that context selection is conditioned on the current context.

Letting $\boldsymbol{z}_1 = (\kappa_1, \boldsymbol{x}_1)$ and $\boldsymbol{z}_2 = (\kappa_2, \boldsymbol{x}_2)$ be two states in $\mathcal{Z}$, we derive the transition probability

$$P_{\boldsymbol{z}_1}(\boldsymbol{z}_2) \triangleq \Pr\left(\boldsymbol{z}(t+1) = (\kappa_2, \boldsymbol{x}_2) | \boldsymbol{z}(t) = (\kappa_1, \boldsymbol{x}_1)\right), \tag{2.3}$$

from $\boldsymbol{z}_1$ to $\boldsymbol{z}_2$ in $\mathcal{Z}$. Note that we can rewrite expression (2.3) as

$$P_{\boldsymbol{z}_1}(\boldsymbol{z}_2) = \Pr\left(\boldsymbol{x}(t+1) = \boldsymbol{x}_2, \kappa(t+1) = \kappa_2 | \boldsymbol{x}(t) = \boldsymbol{x}_1, \kappa(t) = \kappa_1\right). \tag{2.4}$$

Using the Bayes theorem, we get

$$\begin{aligned}
&P_{\boldsymbol{z}_1}(\boldsymbol{z}_2) \\
&\quad = \Pr\left(\boldsymbol{x}(t+1) = \boldsymbol{x}_2 | \kappa(t+1) = \kappa_2, \boldsymbol{x}(t) = \boldsymbol{x}_1, \kappa(t) = \kappa_1\right) \\
&\quad \times \Pr\left(\kappa(t+1) = \kappa_2 | \boldsymbol{x}(t) = \boldsymbol{x}_1, \kappa(t) = \kappa_1\right) \\
&\quad = \Pr\left(\boldsymbol{x}(t+1) = \boldsymbol{x}_2 | \kappa(t+1) = \kappa_2, \boldsymbol{x}(t) = \boldsymbol{x}_1, \kappa(t) = \kappa_1\right) \\
&\quad \times \Pr\left(\kappa(t+1) = \kappa_2 | \kappa(t) = \kappa_1\right) \\
&\quad = \mathbf{1}(\kappa_2 = \kappa_1) \Pr\left(\kappa(t+1) = \kappa_1 | \kappa(t) = \kappa_1\right) \\
&\quad \times \Pr\left(\boldsymbol{x}(t+1) = \boldsymbol{x}_2 | \kappa(t+1) = \kappa(t) = \kappa_1, \boldsymbol{x}(t) = \boldsymbol{x}_1\right) \\
&\quad + \mathbf{1}(\kappa_2 \neq \kappa_1) \Pr\left(\kappa(t+1) = \kappa_2 | \kappa(t) = \kappa_1\right) \\
&\quad \times \Pr\left(\boldsymbol{x}(t+1) = \boldsymbol{x}_2 | \kappa(t+1) = \kappa_2, \boldsymbol{x}(t) = \boldsymbol{x}_1, \kappa(t) = \kappa_1\right),
\end{aligned} \tag{2.5}$$

where $\mathbf{1}(\cdot)$ is the indicator function. Furthermore, we have

$$\Pr\left(\kappa(t+1)=\kappa_1|\kappa(t)=\kappa_1\right)=(1-q)+q\,r_{\kappa_1}, \tag{2.6}$$

and when $\kappa_1 \neq \kappa_2$, we get

$$\Pr\left(\kappa(t+1)=\kappa_2|\kappa(t)=\kappa_1\right)=q\,r_{\kappa_2}. \tag{2.7}$$

Here, q is the probability of switching context, and r_{κ_i} is the probability of selecting context κ_i during a switch.

A transition from GAP $\boldsymbol{x}_1$ to GAP $\boldsymbol{x}_2$ may occur either according to the constituent network at instant $t+1$, or through an appropriate number of random perturbations, but not both.

Let us define F_l by

$$F_l(\boldsymbol{x}_1,\boldsymbol{x}_2) \stackrel{\triangle}{=} \mathbf{1}(\boldsymbol{f}_l(\boldsymbol{x}_1)=\boldsymbol{x}_2). \tag{2.8}$$

Then, we have

$$\begin{aligned}&\Pr\left(\boldsymbol{x}(t+1)=\boldsymbol{x}_2|\kappa(t+1)=\kappa(t)=\kappa_1,\boldsymbol{x}(t)=\boldsymbol{x}_1\right)=\\&\quad\left[(1-p)^n F_{\kappa_2}(\boldsymbol{x}_1,\boldsymbol{x}_2)+(1-p)^{(n-D(\boldsymbol{x}_1,\boldsymbol{x}_2))}p^{D(\boldsymbol{x}_1,\boldsymbol{x}_2)}\mathbf{1}(D(\boldsymbol{x}_1,\boldsymbol{x}_2)\neq 0)\right]\end{aligned} \tag{2.9}$$

and, for $\kappa_1 \neq \kappa_2$, we obtain

$$\begin{aligned}&\Pr\left(\boldsymbol{x}(t+1)=\boldsymbol{x}_2|\kappa(t+1)=\kappa_2,\boldsymbol{x}(t)=\boldsymbol{x}_1,\kappa(t)=\kappa_1\right)=\\&\quad\left[(1-p)^n F_{\kappa_2}(\boldsymbol{x}_1,\boldsymbol{x}_2)+(1-p)^{(n-D(\boldsymbol{x}_1,\boldsymbol{x}_2))}p^{D(\boldsymbol{x}_1,\boldsymbol{x}_2)}\mathbf{1}(D(\boldsymbol{x}_1,\boldsymbol{x}_2)\neq 0)\right]\end{aligned} \tag{2.10}$$

where $D(\boldsymbol{x}_1,\boldsymbol{x}_2)$ is the Hamming distance between two gene-activity profiles $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. The first parts of (2.9) and (2.10) corresponds to the probability of transition from GAP $\boldsymbol{x}_1$ to GAP $\boldsymbol{x}_2$ according to the predictor functions defined by the constituent network at time instant $t+1$. The remaining terms account for transition between GAPs that are due to random gene perturbation.

By replacing the terms of expression (2.5) by their equivalents from (2.6), (2.7), (2.9) and (2.10), it can be shown that the probability of transition from any state $\boldsymbol{z}_1 = (\kappa_1, \boldsymbol{x}_1)$ to any other state $\boldsymbol{z}_2 = (\kappa_2, \boldsymbol{x}_2)$ in $\mathcal{Z}$ is given by

$$P_{\boldsymbol{z}_1}(\boldsymbol{z}_2) = \Big[(1-p)^n \mathbf{1}(\boldsymbol{f}_{\kappa_2}(\boldsymbol{x}_1) = \boldsymbol{x}_2) + (1-p)^{(n-D(\boldsymbol{x}_1,\boldsymbol{x}_2))} p^{D(\boldsymbol{x}_1,\boldsymbol{x}_2)} \mathbf{1}(D(\boldsymbol{x}_1, \boldsymbol{x}_2) \neq 0)\Big] \times \left[\mathbf{1}(\kappa_2 = \kappa_1)((1-q) + q\, r_{\kappa_1}) + \mathbf{1}(\kappa_2 \neq \kappa_1) q\, r_{\kappa_2}\right] \cdot \tag{2.11}$$

Gene perturbation ensures that all the states in the Markov chain of context-sensitive PBN oriented graph communicate; hence, the finite-state Markov chain has a unique steady-state distribution [34].

3. Instantaneously Random Probabilistic Boolean Networks

As we mentioned earlier, The switching frequency of the context differentiates the instantaneously random PBN from the context-sensitive PBN. If the contexts change at every instant, i.e. $q = 1$, then the PBN is instantaneously random [22].

For an instantaneously random PBN, the state of the oriented graph only consists of a GAP. Hence, an instantaneously random PBN with n genes has 2^n states, as opposed to $2^n \times k$ states for a context-sensitive PBN with n genes and k contexts. By the Markovian property, the dynamic behavior of an instantaneously random PBN whose states are in the set of all GAPs $\mathcal{X}$ can be represented as a stationary discrete-time equation

$$\boldsymbol{x}(t+1) = f(\boldsymbol{x}(t), w(t)) \qquad t = 0, 1, \ldots \tag{2.12}$$

where $\boldsymbol{x}(t)$ is the GAP at the updating epoch t. The disturbance $w(t)$ is the manifestation of uncertainties in the biological system, due either to context switching or change in the genes resulting from random gene perturbation.

The transition probability distributions of an instantaneously random PBN is derived in [22]. Here, we can obtain the transition probability matrix of the instantaneously random PBN oriented graph corresponding to a context-sensitive PBN with similar parameters from expression (2.11), when the context is allowed to switch at each epoch by setting $q = 1$. In this case, we can rewrite expression (2.11) as

$$P(\boldsymbol{x}_1, \boldsymbol{x}_2) = (1-p)^n \left[\left(\frac{p}{1-p} \right)^{D(\boldsymbol{x}_1, \boldsymbol{x}_2)} \mathbf{1}\left(D(\boldsymbol{x}_1, \boldsymbol{x}_2) \neq 0 \right) + \sum_{\kappa_2} r_{\kappa_2} \mathbf{1}\left(\boldsymbol{f}_{\kappa_2}(\boldsymbol{x}_1) = \boldsymbol{x}_2 \right) \right] \tag{2.13}$$

for the transition probability distribution between any two states of instantaneously random PBN $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in $\mathcal{X}$.

B. Asynchronous Markovian Regulatory Networks

Two factors motivate the adoption of synchronous updating schemes in Markovian regulatory networks: (1) they are more mathematically tractable; (2) they require significantly less data for inference. In particular, substantial time-course data are required to characterize asynchronism.

On the other hand, the synchronous abstraction is used under the implicit assumption that asynchronous updating will not unduly alter the properties of a system central to the application of interest [35]. Clearly, some properties will be altered. Various potential issues with synchronous networks have been noted. For instance, synchronous abstraction may produce spurious attractors in networks [15]. In the same vein, deviation from synchronous updating modifies the attractor structure of Boolean networks [16] and can change their long-run behavior [36]. From a biological perspective, interactions among genes causing transcription, translation, and degradation occur over a wide range of time-scales [12, 13]. These observations suggest

that we should relax the synchronism assumption in regulatory networks to provide a better description of underlying biological processes.

In this section, we propose two new asynchronous Markovian regulatory networks. The first model introduces asynchronism in regulatory networks at the gene level. The second model extends this approach by considering asynchronism at the state level. Whereas the first method is akin to currently proposed asynchronous models, we will argue that the second approach is more suitable from a translational perspective.

The first proposed asynchronous model introduces asynchronism into Markovian regulatory networks by updating each gene based on its period. This approach of introducing asynchronism into networks extends the currently favored approach of studying asynchronism in regulatory models. The proposed model, called a *deterministic-asynchronous probabilistic Boolean network* (DA-PBN), is an extension of probabilistic Boolean networks in which the time scales of biological updates at various genes can be different.

A DA-PBN is an extension of PBN in which different time-scales for various biological processes are allowed. Each gene, or biological component, is updated based on an individual period, which may differ from one component to another. Yet, the updating period of each gene is fixed given the context of the network. The intent of context-sensitivity is to incorporate the effects of latent variables not directly captured in the model. The behavior of these latent variables influences both regulation and updating periods of genes. The uncertainty about the context of a regulatory network resulting from latent variables is captured through a probability measure on the possible contexts. The exact updating periods and functions of genes cannot be practically specified. At best, we can estimate the set of possible updating periods and corresponding updating functions for each gene. As a stochastic Boolean

network with asynchronous updates, a DA-PBN expands the benefits of traditional PBNs by adding the ability to cope with temporal context as well as regulatory context.

Our objectives for introducing asynchronism via the DA-PBN model are twofold. First, we show that the synchronous formalism of PBNs can be relaxed to introduce asynchronous PBNs. Second, we provide a methodology to derive optimal intervention strategies for the DA-PBN model, which will be discussed in Chapter V.

The delay and the updating order of a given gene are only observable with respect to the activity levels of other genes and proteins involved in the regulation process [14]. Thus, it is impractical to study the alteration of one specific gene over time, while keeping the levels of all other genes in the model constant. Practically, we can measure the gene-expression profile (or states) at each observation instant. The inter-transition interval between two states can then be modeled by a random variable.

In [29] and [30], experimentally validated Boolean rules are considered. Under a synchronous assumption, the oriented graphs can accurately determine the phenotypic behavior of the underlying biological processes. However, these studies suggest that asynchronously updating the genes when utilizing the same Boolean rules generates very complex pathways which possess many incompatible and unrealistic phenotypes. Although not mentioned explicitly in [29] and [30], it is evident that asynchronously updating the genes changes the global behavior of regulatory networks by changing their oriented graphs.

The results of [29] and [30] indicate that rule-based regulatory models should maintain the topology of the oriented graph generated by experimentally validated predictor rules, as if the genes are coupled. In other words, regulatory models should accurately translate the logical relationships, i.e. the regulatory graph, governing the inter-gene interactions into the oriented graph specifying the dynamics of the model.

Moreover, they should enable the analysis of the temporal behavior of biological systems. Since our objective is to alter the long-run behavior of biological systems via effective intervention strategies, not only should our regulatory models possess the previous two attributes, but these models should also be inferable from empirical data.

Having these in mind, we propose another asynchronous regulatory network model: *semi-Markov asynchronous regulatory networks* (SM-ARN). In the SM-ARN, the asynchronism is at the state level. In this model, the empirically measurable timing information of biological systems is incorporated into the model. This temporal information determines the typical time delay between transitions from one state to another. Since the order of updating genes and their relative time delays depend on the levels of other regulatory components, estimating the updating time of each gene in isolation, and independent of the values of other genes, is highly problematic, if not impossible. Time-course data enable the estimation of inter-transition times between states, not the updating time of each gene. It is then natural to introduce asynchronism at the state level.

An SM-ARN is specified with two sets of information. The first set determines the rule-based multivariate interactions between genes. Considering simultaneous updating, we can specify the oriented graph of the model based on this information. In other words, the first set of information specifies a PBN, which is generated from a given set of Boolean functions for updating each gene. The generated oriented graph guarantees the predictability of the rule-based topology. The second set of information consists of the distributions of inter-transition intervals between any two states that are directly connected. These values can be empirically inferred from time-course data.

1. Deterministic-Asynchronous Probabilistic Boolean Networks

To date, asynchronism has been introduced into Boolean networks by updating each gene based on its period. These studies try to understand generic characteristics of asynchronous updating schemes in randomly generated Boolean networks. To accomplish this aim, a wide range of artificial asynchronous updating protocols with different degrees of freedom in the selection of the updating period for each gene has been postulated.

We categorize previously proposed asynchronous protocols into two groups. In the first category, termed *stochastic asynchronous* protocols, the updating period of each gene is randomly selected based on a given distribution [16,36,37]. In the second category, termed *deterministic asynchronous* protocols, the updating period of each gene is fixed, but can differ from one gene to another [35, 36, 38]. There have also been studies that consider both stochastic and deterministic asynchronous protocols in an effort to investigate the predictability of Boolean networks when asynchronous updating schemes are used instead of synchronous ones [29, 39].

The study of both randomly generated and experimentally validated Boolean networks reveals that stochastic asynchronism has some limitations. Stochastic asynchronous updating methods can significantly change the properties of oriented graphs [35,38]. Starting from wild-type gene expressions, neither the Boolean networks of [39] or [29] successfully predict the anticipated long-run attractors of their networks. Earlier studies indicate that constraining the degrees of freedom in the asynchronous protocols can improve the predictability of Boolean networks. More structured asynchronous protocols predict the long-run behavior of Boolean networks more effectively by representing their cyclic attractors [29,37]. It must be recalled that in all of these studies of asynchronism the timing protocols have been modeled mathematically with-

out biological verification. At this point, perhaps all we can say is that synchronism or asynchronism are modeling assumptions and the choice in a specific circumstance depends upon the available data and application.

When the context of a biological system is known, there is a consensus that asynchronism in regulatory networks is deterministic rather than random [35]; however, deterministic asynchronous Boolean networks pose practical challenges. Even if we can measure the level of each gene in isolation while the other genes remain constant, at best, we could produce estimates for updating periods. Owing to the effects of measurement noise and the existence of latent variables, we cannot exactly specify them.

Focusing on the effects of latent variables, as is customary when considering probabilistic networks, at best we can estimate a set consisting of the most probable updating periods for each gene in the network; each set depending on the status of latent variables. A set of updating periods, whose members are the deterministic periods of each gene in the regulatory network, defines the updating protocol of a *deterministic-asynchronous Boolean network*. This means that there is a finite collection of deterministic asynchronous Boolean networks that defines the dynamics of the system. The updating periods of genes depend on the temporal context of the biological system, which can be influenced by latent variables.

Having the probabilities of selecting each context, the model selects one of the constituent deterministic asynchronous Boolean networks at each updating instant. The system evolves according to the selected constituent deterministic asynchronous Boolean network until its constituent network changes. This approach of introducing asynchronism into PBNs extends the currently favored approach of studying asynchronism in regulatory models. The proposed model, called a *deterministic-asynchronous probabilistic Boolean network*, is an extension of probabilistic Boolean networks in

which the time scales of various biological updates can be different. The term "probabilistic" emphasizes the random selection of a context, while the term "deterministic" refers to the asynchronous protocol within each context of the regulatory network.

For consistency with context-sensitive PBN, we use the same notation to define DA-PBN. As with a synchronous PBN, in a DA-PBN, gene values are quantized to two levels. A DA-PBN consists of a sequence $V = \{x_i\}_{i=1}^{n}$, of n nodes, where $x_i \in \{0, 1\}$. Each x_i represents the expression value of a gene, which is either ON or OFF. A DA-PBN is composed of a collection of k constituent deterministic-asynchronous Boolean networks (DA-BNs). In a DA-PBN, the active DA-BN changes at updating instants selected by a binary switching random variable. A DA-PBN acts like one of its constituent DA-BNs, each being referred to as a *context*, between two switching instants.

The l-th DA-BN $(V, \boldsymbol{f}_l, \boldsymbol{\Theta}_l)$ is defined by two vector-valued functions. The vector-valued function $\boldsymbol{f}_l$ consists of n predictors, $\boldsymbol{f}_l = (f_{l1}, \ldots, f_{ln})$, where $f_{li} : \{0, 1\}^n \to \{0, 1\}$ denotes the predictor of gene i, whenever context l is selected. The vector-valued function $\boldsymbol{\Theta}_l$ consists of n updating components, $\boldsymbol{\Theta}_l = (\theta_{l1}, \ldots, \theta_{ln})$. Each function $\theta_{li} : \mathbb{N} \to \{0, 1\}$ is defined with a pair of fixed parameters, (a_{li}, b_{li}). The parameter $a_{li} \in \mathbb{N}$ specifies the updating period of gene i, when context l is selected. The parameter $b_{li} \in \{0, \ldots, a_{li}-1\}$ further differentiates the exact updating instant of each gene within its updating period. The two degrees of freedom in θ_{li} are sufficient to assign any instant of time $t \in \mathbb{N}$ as the updating epoch of gene i:

$$\theta_{li}(t) \triangleq \begin{cases} 1, & \text{if} \quad t \equiv b_{li} \pmod{a_{li}} \\ 0, & \text{otherwise.} \end{cases} \tag{2.14}$$

Similar to context-sensitive PBN, we focus on DA-PBN with perturbations, in which each gene may change its value with small probability p at each epoch.

At each updating instant a decision is made whether to switch the current constituent DA-BN. The switching probability q is a system parameter. If the current DA-BN is not switched, then the DA-PBN behaves as a fixed DA-BN and genes are updated synchronously according to the current constituent network:

$$x_i(t+1) = \begin{cases} \mathbf{1}(\boldsymbol{\gamma}(t+1) \neq \mathbf{0}) \left(x_i(t) \oplus \gamma_i(t+1)\right) & \\ \qquad +\mathbf{1}(\boldsymbol{\gamma}(t+1) = \mathbf{0}) f_{li}\left(x_1(t), \ldots, x_n(t)\right), & \text{if } \theta_{li}(t+1) = 1 \\ x_i(t), & \text{if } \theta_{li}(t+1) = 0 \end{cases} \tag{2.15}$$

where $\gamma_i(t)$ is a Bernoulli random variable with parameter p and the random vector $\boldsymbol{\gamma}$ at instant t is defined as $\boldsymbol{\gamma}(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t))$. The operator $\oplus$ is component-wise addition in modulo two and f_{li} is the predictor of gene i according to the DA-BN l.

If a switch occurs, then a new constituent network is randomly selected according to a selection probability measure $\{r_l\}_{l=1}^{k}$. After selecting the new constituent network $\hat{l}$, the values of the genes are updated using (2.15), but with $\boldsymbol{f}_{\hat{l}}$ and $\boldsymbol{\Theta}_{\hat{l}}$ instead.

2. Semi-Markov Asynchronous Regulatory Networks

Although DA-PBN can be used to model biological systems, earlier research suggests that the assumption of asynchronism at the gene level has a number of drawbacks. Asynchronously updating the genes changes the global behavior of regulatory networks due to changing their oriented graph, which models the dynamics of the system [14]. Along this line, it has been shown that small perturbations do not settle down in a random Boolean network with gene-level asynchronism. Consequently, the asynchronous network is in the chaotic regime while its synchronous counterpart is in the critical regime [40]. The studies of experimentally validated Boolean networks in [29] and [30] suggest that oriented graphs of given Boolean networks provide

accurate predictability, whereas the oriented graphs of networks utilizing the same Boolean rules with asynchronously updated genes are very complex and possess many incompatible or unrealistic pathways.

From these observations, we gather that an asynchronous regulatory model should maintain the topology of the oriented graph as specified by the logical rules governing the interactions between genes. In other words, regulatory models should accurately translate the logical relationships, i.e. the regulatory graph, governing the interactions of genes to the oriented graph specifying the dynamics of the model. Moreover, they should enable the analysis of the temporal behaviors of biological systems.

Due to the aforementioned observations, we propose *semi-Markov asynchronous regulatory networks*. For consistency with previous networks, we use the same notation to define SM-ARNs. We consider a sequence of n genes, $V = \{x_i\}_{i=1}^{n}$, representing the expression values of genes involved in the model. The expression value of gene i at time t $x_i(t)$ is selected from two quantization levels. The states of an SM-ARN are defined as the ordered pairs consisting of a constituent network κ and a gene-activity profile $\boldsymbol{x}$. As in Section II.A, the GAP can be considered to be an n-tuple vector $\boldsymbol{x}(t) = (x_1(t), \ldots, x_n(t))$ giving the values of genes at time t, where $\boldsymbol{x}(t) \in \mathcal{X} = \{0,1\}^n$. At each time $t \in \mathbb{R}^+$, the state of SM-ARN $\boldsymbol{z}(t)$ is selected from the set of all possible states

$$\mathcal{Z} = \{(\kappa, \boldsymbol{x}) : \kappa \in \{1, \ldots, k\}, \boldsymbol{x} \in \mathcal{X}\} \cdot \tag{2.16}$$

Considering two consecutive epochs t_k and t_{k+1} per Figure 4, the state of the SM-ARN for all the times $t_k \leq t < t_{k+1}$ is $\boldsymbol{z}(t_k) = \boldsymbol{z}_k$. At t_{k+1}, the model enters a new state $\boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_{k+1}$. If τ_{k+1} is the time spent in state $\boldsymbol{z}_k$ prior to transition to state $\boldsymbol{z}_{k+1}$, then we have $\tau_{k+1} = t_{k+1} - t_k$. In the SM-ARN model, this inter-transition

interval is modeled with a non-negative random variable with probability distribution

$$P_{\boldsymbol{z}_k\boldsymbol{z}_{k+1}}(\tau) \triangleq \Pr\left(\tau_{k+1} \leq \tau | \boldsymbol{z}(t_k) = \boldsymbol{z}_k, \boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_{k+1}\right), \tag{2.17}$$

where $\boldsymbol{z}_k$ and $\boldsymbol{z}_{k+1}$ are in $\mathcal{Z}$. According to (2.17), the probability distribution of sojourn time in the current state $\boldsymbol{z}_k$ prior to transition to the successive state $\boldsymbol{z}_{k+1}$ could depend on both states. We require the inter-transition interval distributions, $P_{\boldsymbol{z}_k\boldsymbol{z}_{k+1}}(\tau)$, for any two directly connected states as one of the two sets of information needed to define an SM-ARN. Time-course data could provide the information leading to these distributions.

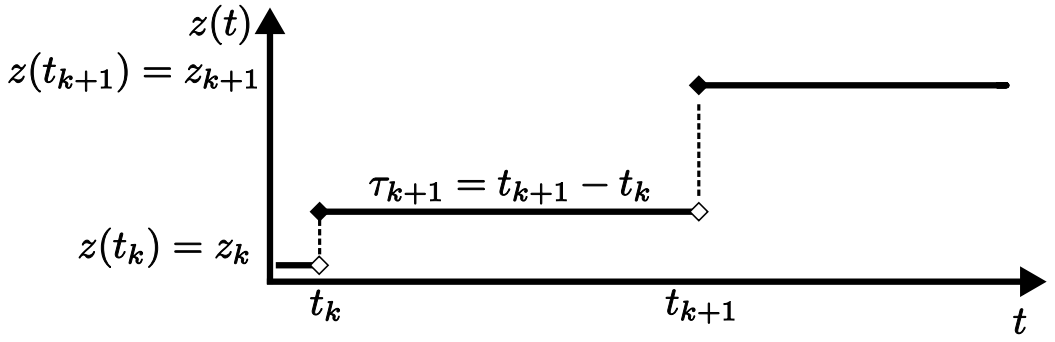


Fig. 4. A schematic of transition in SM-ARN with two consecutive epoch times t_k and t_{k+1}. The inter-transition interval, τ_{k+1}, is the sojourn time in state $\boldsymbol{z}_k$ prior to the transition to state $\boldsymbol{z}_{k+1}$.

In the SM-ARN model, the empirically measurable timing information of biological systems is incorporated into the model. This measurable timing information determines the typical time delay between transitions from state $\boldsymbol{z}_k$ to state $\boldsymbol{z}_{k+1}$. Since the order of updating genes and their relative time delays depend on the levels of other regulatory components, we are able to estimate the joint distribution of updating time of vectored value of genes, and not their marginal distributions. The estimation of the latter requires measurements providing information on the updating time of each gene in isolation, and independent of the values of other genes.

Design of experiments that provide this temporal information is highly problematic, if not impossible. Time-course data enable the estimation of inter-transition times between states in $\mathcal{Z}$, not the updating time of each gene. It is at the state level that asynchronism is introduced in an SM-ARN, and not the gene level.

Borrowing the methodology proposed in Section II.A, we proceed to define the *embedded-PBN* of an SM-ARN. The embedded-PBN of an SM-ARN models the probabilistic rule-based connections of gene interactions and constitutes the other set of information required for specification of an SM-ARN. The embedded-PBN specifies the oriented graph of the SM-ARN based on the predictors of the genes. The oriented graph of an SM-ARN is a directed graph whose vertices are the states of the SM-ARN in $\mathcal{Z}$, and for which there is an edge between any two directly connected states. The weight of each edge is the transition probability between two vertices connected by that edge.

Let $\{\boldsymbol{f}_l\}_{l=1}^{k}$ be the set of k realizations of the embedded-PBN. If the genes are coupled, then at each simultaneously updating instant, one of the k possible realizations of the embedded-PBN is selected. Each vector-valued function $\boldsymbol{f}_l$ has the form $\boldsymbol{f}_l = (f_{l1}, \ldots, f_{ln})$. Each function $f_{li} : \{0,1\}^n \to \{0,1\}$ denotes the predictor of gene i, when the realization l is selected. At each simultaneous updating instant a decision is made whether to switch the context of the network. If at a particular updating instant, it is decided that the realization of the network should not be switched, then the embedded-PBN behaves as a fixed Boolean network and simultaneously updates the values of all the genes according to their current predictors. If it is decided that the network should be switched, then a realization of the embedded-PBN is randomly selected according to a selection distribution $\{r_l\}_{l=1}^{k}$. After selecting the vector-valued function $\boldsymbol{f}_l$, the values of the genes are updated according to the predictors determined by $\mathbf{f}_l$. We assume that the probability of selecting the ith realization of the

embedded-PBN r_i is known. In addition, we allow perturbations in the embedded-PBN, whereby each gene may change its value with a small probability p at each updating instant.

The graph specifying the relationships among the states of an embedded-PBN, defined as above, can be represented as a Markov chain whose transition probability is given by (2.11). On the other hand, the graph of the relationships among the states specified by the embedded-PBN is the regulatory graph of the SM-ARN. Originating from a state $\boldsymbol{z}(t_k) = \boldsymbol{z}_k$, the successor state $\boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_{k+1}$ is selected randomly within the set $\mathcal{Z}$ according to the transition probability $p_{\boldsymbol{z}_k}(\boldsymbol{z}_{k+1})$ defined by regulatory graph of the SM-ARN:

$$p_{\boldsymbol{z}_k}(\boldsymbol{z}_{k+1}) \triangleq \Pr\left(\boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_{k+1} | \boldsymbol{z}(t_k) = \boldsymbol{z}_k\right), \tag{2.18}$$

for all $\boldsymbol{z}_k, \boldsymbol{z}_{k+1} \in \mathcal{Z}$. In other words, the oriented graph of an SM-ARN is the same as its regulatory graph. However, the update of states in the oriented graph of an SM-ARN occurs on various time-scales according to inter-transition interval distributions. Therefore, the oriented graph of the SM-ARN defined by the embedded-PBN maintains the topology of the oriented graph generated by the experimentally validated predictors of genes.

Gene perturbation ensures that all the states of a SM-ARN communicate in its oriented graph. Hence, the fraction of time that the SM-ARN spends in state $\boldsymbol{z} \in \mathcal{Z}$ in the long run can be computed [41]:

$$p(\boldsymbol{z}) = \frac{\pi(\boldsymbol{z})\overline{\tau}(\boldsymbol{z})}{\sum_{\boldsymbol{z}\in\mathcal{Z}} \pi(\boldsymbol{z})\overline{\tau}(\boldsymbol{z})}, \tag{2.19}$$

with probability one. Here, $\{\pi(\boldsymbol{z})\}_{\boldsymbol{z}\in\mathcal{Z}}$ is the steady-state distribution of the states in the Markov chain representing the oriented graph of the SM-ARN, and $\{\overline{\tau}(\boldsymbol{z})\}_{\boldsymbol{z}\in\mathcal{Z}}$ is

the expected sojourn time in state $\boldsymbol{z}$, which can be computed from the information in (2.18) and (2.17). One can easily verify that $p(\boldsymbol{z})$, the fraction of time spent in state $\boldsymbol{z}$ in the long run, will be equal to $\pi(\boldsymbol{z})$, which is the fraction of the transitions to state $\boldsymbol{z}$ if all the nodes are synchronously updated.

CHAPTER III

MARKOVIAN REGULATORY MODELS: CASE STUDIES *

In this chapter, we present two case studies which are used throughout this volume. We construct four regulatory models based on these case studies. First, we construct PBNs using gene-expression data collected in a profiling study of metastatic melanoma [31, 42]. We utilize the profiling study of metastatic melanoma to infer a 10-gene instantaneously random PBN as well as a 7-gene context-sensitive PBN. Second, a context-sensitive PBN is obtained by suitably extending a Boolean model proposed in [29] for modeling mammalian cell cycle regulation in Section III.B. This network postulates a generic mammalian cell cycle with a mutated phenotype [14]. At last in Section III.C, an SM-ARN is devised using the regulatory model in Section III.B as its embedded-PBN.

A. Metastatic Melanoma Gene Expression: Probabilistic Boolean Network

In this section, we construct a 10-gene instantaneously random PBN and a 7-gene context-sensitive PBN based on data collected in a profiling study of metastatic melanoma in which the abundance of messenger RNA for the gene $Wnt5a$ was found to be highly discriminating between cells with properties typically associated with

high metastatic competence versus those with low metastatic competence [8].

These findings were validated and expanded in a second study, in which experimentally increasing the levels of the $Wnt5a$ protein secreted by a melanoma cell line via genetic engineering methods directly altered the metastatic competence of that cell as measured by the standard in vitro assays for metastasis [43]. A further finding of interest in this study is that an intervention that blockes the $Wnt5a$ protein from activating its receptor, with the help of an antibody that binds the $Wnt5a$ protein, can substantially reduce $Wnt5a$'s ability to induce a metastatic phenotype. This suggests intervention based on a strategy that alters the contribution of the $Wnt5a$ gene to biological regulation. Disruption of this influence can potentially reduce the chance of a melanoma metastasizing, a desirable outcome.

Ten genes, including the $Wnt5a$ gene, were selected in [25] based on the predictive relationships among 587 genes: $Wnt5a$, *pirin*, $S100p$, $Ret1$, $Mmp3$, *Phoc*, $Mart1$, *Hadhb*, *Synuclein*, and $Stc3$. We apply a modification of the design procedure proposed in [44] to generate a 10-gene instantaneously random PBN possessing four constituent BNs [42]. The method of [44] generates BNs with given attractor structures and the overall PBN is designed so that the data points, which are assumed to come from the steady-state behavior of the network, are attractors in the designed PBN. However, this method doesn't consider the distribution of attractors in the empirical data. Considering the available data, we have to modify this method to arrive at a PBN resembling a cancerous situation. The modified method searches in the set of possible Boolean networks generated by the method of [44], and finds Boolean networks whose states with up-regulated $Wnt5a$ possess larger aggregated probability than the states with down-regulated $Wnt5a$. The attractors with up-regulated $Wnt5a$ are only %16 of all the observed attractors in the metastatic melanoma dataset in [42].

The regulatory graphs of these 10-gene BNs are given in Figure 5. In binary representation of the GAPs, the order of the genes is as listed earlier with $Wnt5a$ being the most significant bit and $Stc3$ being the least significant bit.

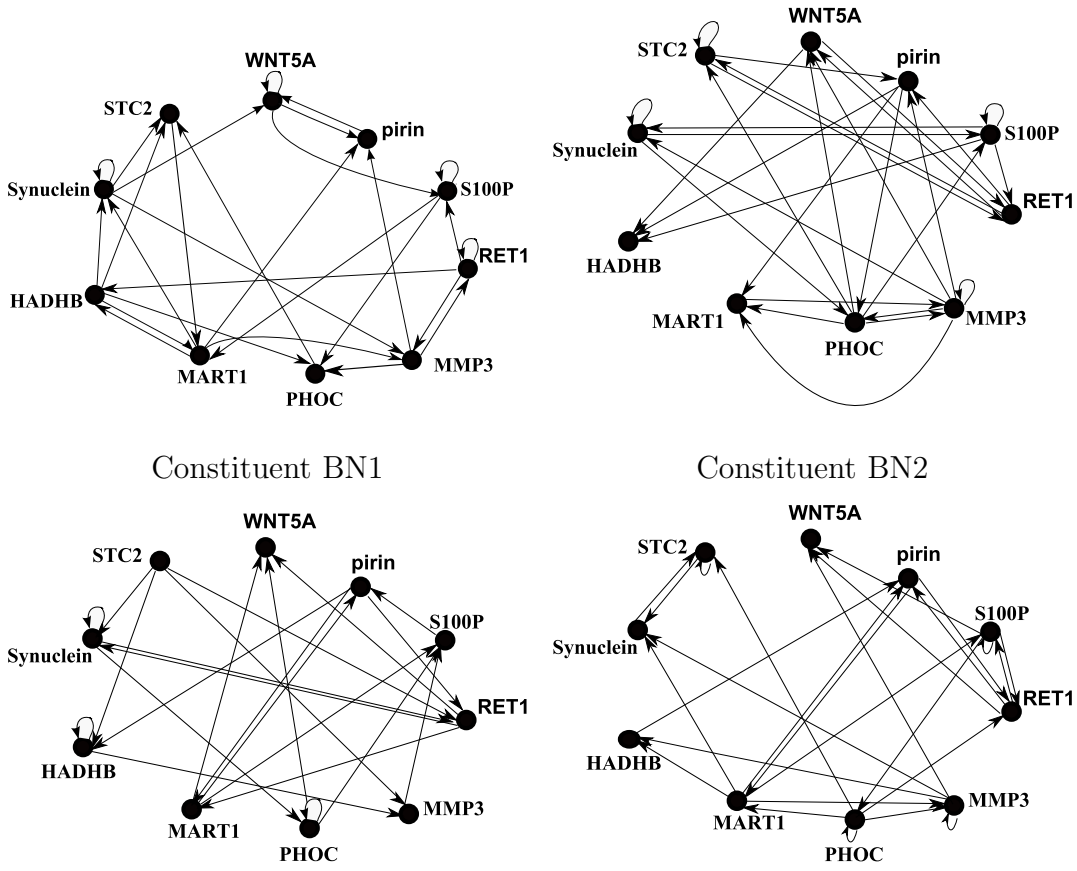


Fig. 5. The regulatory graphs of the four constituent Boolean networks used to construct a 10-gene instantaneously random PBN for the metastatic melanoma data.

We also inferred a 7-gene context-sensitive PBN from the same profiling study of metastatic melanoma by considering genes: $Wnt5a$, *pirin*, $S100p$, $Ret1$, $Mart1$,

Hadhb, and *Stc*3. Computational limitations forced us to consider the 7-gene context-sensitive PBN with 512 states, instead of 10-gene networks with 4096 states. We use the same procedure explained above to infer the constituent Boolean networks. To generate the context-sensitive PBN based on the inferred Boolean networks, we set both the switching and perturbation probabilities to 0.01. The selection probability distribution is assumed to be uniform $\{r_l = 0.25\}_{l=1}^{4}$. The constituent networks $\{\boldsymbol{f}_l\}_{l=1}^{4}$ are reported in tables I, II, III and IV, respectively.

Each of Table I to Table IV has $2^{pred} + n$ rows and n columns. Variable $pred$ denotes the maximum number of predictors for each of the n genes in the network. We set $pred = 3$ in this study. The top 2^{pred} rows depicts the predictor functions of the genes. We separate the top part of each table from its lower part with a horizontal line to increase the readability. The lower n rows of each table provide the predictors for the genes in the Boolean network. For example, genes 3, 5, and 7 are the predictors of gene 1 in the constituent network $\boldsymbol{f}_l$ according to the 9-th row of Table I. Hence, $f_{11}(x_3, x_5, x_7)$, the predictor function of gene 1, can be specified by its 8 possible outcomes enumerated in the first column of Table I. Whenever the number of predictors is less than $pred = 3$ the outcomes of the predictor function can be enumerated with less than 2^{pred} values. For instance, gene 3 in Table I has two predictors (refer to row $2^3 + 3$ of Table I), so its predictor function $f_{13}(x_3, x_1)$ can be fully specified with 4 values. According to the the upper part of the third column of Table I, the value of gene 3 is set to 0 when the values of genes $x_3(t)$ and $x_1(t)$ are 0 and 0, respectively.

The intervention objective for these two networks is to down-regulate the *Wnt*5*a*, because this gene ceasing to be down-regulated is strongly predictive of the onset of metastasis. We will present two intervention methods using the constructed PBNs, with the aim of down-regulating the *Wnt*5*a* gene. These models are used because the

Table I. Constituent network $\boldsymbol{f}_1$.

1	1	0	1	0	0	1
1	0	0	0	0	0	0
1	1	1	1	1	0	1
1	1	1	1	1	0	1
1	1		0		1	0
0	0		0		1	1
0	1		0		1	1
0	1		0		1	0

3	5	7
2	6	1
3	1	
2	4	7
3	7	
5	7	1
3	7	1

Table II. Constituent network $\boldsymbol{f}_2$.

0	0	0	1	1	0	1
1	0	0	1	0	0	1
0	1	1	1	1	1	0
1	1	1	1	1	1	1
1			0		0	
0			1		1	
0			0		1	
0			1		1	

2	6	1
2	6	
2	5	
2	4	7
3	4	
2	5	
5	7	

Table III. Constituent network $\boldsymbol{f}_3$.

1	0	0	0	1	0	1
1	1	0	0	0	0	0
1	1	1	0	0	1	0
0	0	0	0	0	1	1
1		1	1	1	1	
1		0	1	1	1	
0		0	0	1	0	
1		0	1	1	1	

4	5	6
4	5	
2	4	1
4	7	1
3	7	1
3	5	6
4	6	

Table IV. Constituent network $\boldsymbol{f}_4$.

1	1	0	0	1	1	1
1	0	0	1	0	0	1
1	1	1	1	1	1	0
1	0	0	1	1	1	1
0	1	0		0	0	
0	1	0		0	1	
1	1	0		0	1	
0	1	1		1	1	

2	5	6
2	4	7
2	5	7
2	6	
3	6	7
3	6	1
6	7	

relation of $Wnt5a$ to metastasis is well established and the binary nature of the up or down regulation suits a binary model.

The first step in devising any therapeutic intervention is to designate desirable and undesirable states. This depends upon the existence of relevant biological knowledge. In these particular examples, our prior knowledge indicates that the status of $Wnt5a$ relates to metastasis in melanoma tumors. A state is desirable if $Wnt5a$ is 0 and undesirable if $Wnt5a$ is equal to 1.

B. Mutated Mammalian Cell Cycle: Context-Sensitive Probabilistic Boolean Network

We consider a context-sensitive PBN that is a probabilistic generalization of the Boolean model proposed in [29] for mammalian cell-cycle regulation. This context-sensitive PBN postulates the mammalian cell cycle with a mutated phenotype. Mutated cells grow in the absence of extra-cellular growth factors [14].

During the late 1970s and early 1980s, yeast geneticists identified the cell-cycle genes encoding for new classes of molecules, including the cyclins (so-called because of their cyclic pattern of activation) and their cyclin dependent kinase (cdk) partners [29]. Our model is rooted in the work of Faure *et al.*, who have recently derived and analyzed the Boolean functions of the mammalian cell cycle [29]. Using these Boolean functions, the authors have been able to quantitatively reproduce the main known features of the wild-type biological system, as well as the consequences of several types of mutations.

Mammalian cell division is tightly controlled. In a growing mammal, the cell division should coordinate with the overall growth of the organism. This coordination is controlled via extra-cellular signals. These signals indicate whether a cell should

divide or remain in a resting state. The positive signals, or growth factors, instigate the activation of Cyclin D ($CycD$) in the cell.

The key genes in this model are $CycD$, retinoblastoma (Rb), and $p27$. Rb is a tumor-suppressor gene. This gene is expressed in the absence of the cyclins, which inhibit Rb by phosphorylation. Whenever $p27$ is present, Rb can be expressed even in the presence of $CycE$ or $CycA$. Gene $p27$ is active in the absence of the cyclins. Whenever $p27$ is present, it blocks the action of $CycE$ or $CycA$. Hence, it arrests the cell cycle. Table V summarizes the Boolean functions of the wild-type cell cycle network, and Figure 6 depicts its regulatory graph.

Table V. Wild-type Boolean functions of mammalian cell cycle.

Product	Predictors
$CycD$	Input
Rb	$(\overline{CycD} \wedge \overline{CycE} \wedge \overline{CycA} \wedge \overline{CycB}) \vee (p27 \wedge \overline{CycD} \wedge \overline{CycB})$
$E2F$	$(\overline{Rb} \wedge \overline{CycA} \wedge \overline{CycB}) \vee (p27 \wedge \overline{Rb} \wedge \overline{CycB})$
$CycE$	$(E2F \wedge \overline{Rb})$
$CycA$	$(E2F \wedge \overline{Rb} \wedge \overline{Cdc20} \wedge (\overline{Cdh1 \wedge UbcH10})) \vee (CycA \wedge \overline{Rb} \wedge \overline{Cdc20} \wedge (\overline{Cdh1 \wedge UbcH10}))$
$p27$	$(\overline{CycD} \wedge \overline{CycE} \wedge \overline{CycA} \wedge \overline{CycB}) \vee (p27 \wedge (\overline{CycE \wedge CycA}) \wedge \overline{CycB} \wedge \overline{CycD})$
$Cdc20$	$CycB$
$Cdh1$	$(\overline{CycA} \wedge \overline{CycB}) \vee (Cdc20)$
$UbcH10$	$(\overline{Cdh1}) \vee (Cdh1 \wedge UbcH10 \wedge (Cdc20 \vee CycA \vee CycB))$
$CycB$	$(\overline{Cdc20} \wedge \overline{Cdh1})$

The preceding explanation represents the wild-type cell-cycle model. Following one of the proposed mutations in [29], we assume that $p27$ is mutated and its logical rule is always zero (OFF). In this cancerous scenario, $p27$ can never be activated. As we mentioned earlier, whenever $p27$ is present, Rb can be expressed even in the presence of $CycE$ or $CycA$. For the mutated cell-cycle network, however, $p27$ is always zero and Rb cannot be expressed in the case where $CycD$ is not present but $CycE$ or

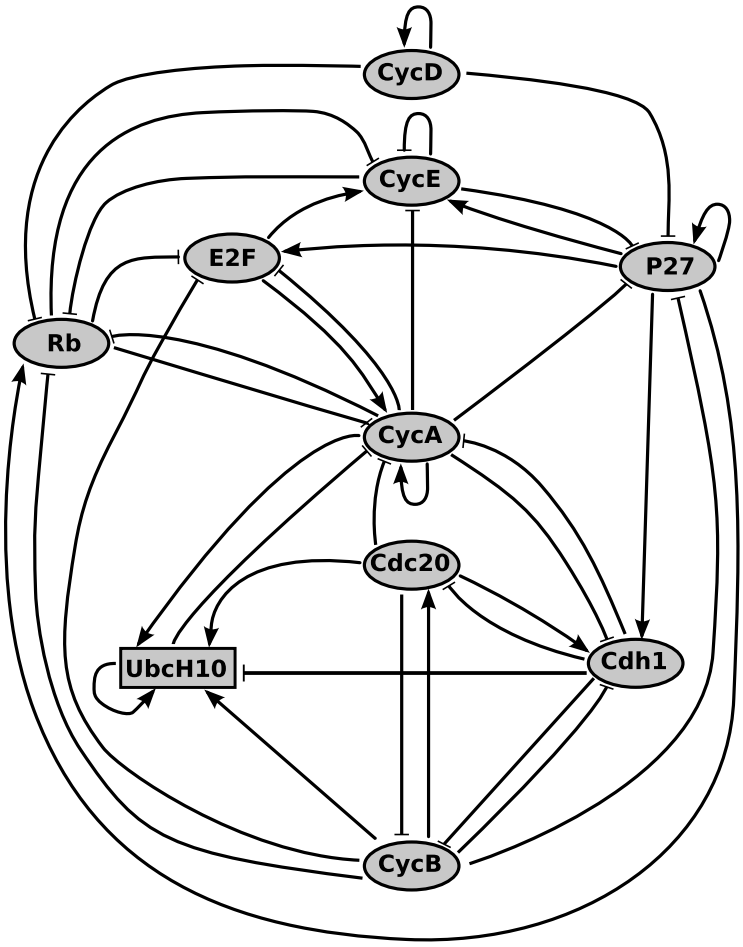


Fig. 6. The regulatory graph of the mammalian cell cycle network as it was presented in Faure *et al.* Each node represents the activity of a key regulatory element. Blunt arrows stand for inhibitory effects, normal arrows for activations.

CycA is active. This mutation introduces a situation where both *CycD* and *Rb* might be inactive. As a result, in this mutated phenotype, the cell cycle continues in the absence of any growth factor. In other words, we consider the states in which both *Rb* and *CycD* are down-regulated as undesirable states, when $p27$ is mutated. Table VI summarizes the mutated Boolean functions, governing the interactions among components in the regulatory graph in Figure 7.

The Boolean functions in Table VI are used to construct a context-sensitive PBN

Table VI. Mutated Boolean functions of mammalian cell cycle.

Product	Predictors
$CycD$	Input
Rb	$(\overline{CycD} \wedge \overline{CycE} \wedge \overline{CycA} \wedge \overline{CycB})$
$E2F$	$(\overline{Rb} \wedge \overline{CycA} \wedge \overline{CycB})$
$CycE$	$(E2F \wedge \overline{Rb})$
$CycA$	$(E2F \wedge \overline{Rb} \wedge \overline{Cdc20} \wedge (\overline{Cdh1} \wedge \overline{UbcH10})) \vee (CycA \wedge \overline{Rb} \wedge \overline{Cdc20} \wedge (\overline{Cdh1} \wedge \overline{UbcH10}))$
$Cdc20$	$CycB$
$Cdh1$	$(\overline{CycA} \wedge \overline{CycB}) \vee (Cdc20)$
$UbcH10$	$(\overline{Cdh1}) \vee (Cdh1 \wedge UbcH10 \wedge (Cdc20 \vee CycA \vee CycB))$
$CycB$	$(\overline{Cdc20} \wedge \overline{Cdh1})$

model for the cell cycle [14]. The extra-cellular signal to the cell-cycle model is considered to be a latent variable. The growth factor is not part of the cell and its value is determined by the surrounding cells. The expression of *CycD* changes independently of the cell's content and reflects the state of the growth factor. Depending on the expression status of *CycD*, we obtain two constituent Boolean networks for the PBN. As shown in Figure 8, the first constituent Boolean network is determined from Table VI when the value of *CycD* is equal to zero. Similarly, the second constituent Boolean network is determined by setting the value of *CycD* to one according to Figure 8. To completely define the context-sensitive PBN, the probability of switching the context, the probability that a gene perturbs, and the probability distribution of selecting each constituent network have to be specified. We assume that these are known. Here, we set the switching and the perturbation probabilities each equal to 0.001, and assume that the two constituent networks are equally likely.

According to Table VI, the mutated cell-cycle PBN consists of nine genes: *CycD*, *Rb*, *E2F*, *CycE*, *CycA*, *Cdc*20, *Cdh*1, UbcH10, and *CycB*. The above order of genes is used in the binary representation of the states of the context-sensitive PBN, with the

context of the networks, *CycD*, as the most significant bit. *Rb* is in the most significant position in the gene-activity profiles and *CycB* is placed at the least significant bit.

Choosing *CycD* and *Rb* as the most significant bits in the state representation of the context-sensitive PBN facilitates the characterization of the undesirable states. We assume that the simultaneous down-regulation of *CycD* and *Rb*, i.e. the cell growth in the absence of growth factors, is undesirable. Consequently, the state-space is readily partitioned into undesirable and desirable states. As mentioned earlier, application of any system-based therapeutic method with external control requires the designation of desirable and undesirable states, and this depends upon the existence of relevant biological knowledge. In the cell-cycle example when $p27$ is mutated, the functionality of the network suggests that the states in which both *Rb* and *CycD* are down-regulated should be avoided. Hence, in the binary representation of states in $\mathcal{Z}$, a state is undesirable if its two most significant positions have value 1, otherwise it is a desirable state.

C. Mutated Mammalian Cell Cycle: Semi-Markov Asynchronous Regulatory Network

In this section, we present an SM-ARN that is devised using the context-sensitive PBN of mammalian cell cycle in Section III.B. The constructed SM-ARN postulates the cell-cycles with mutated phenotype in which cell cycle cannot be arrested in the absence of extra-cellular growth factors.

The Boolean functions in Table VI are used to construct the embedded-PBN of the cell-cycle's SM-ARN. The defined embedded-PBN maintains the topology of the oriented graph generated by the predictors in Table VI. To this end, we assume that the extra-cellular signal to the cell-cycle model is a latent variable. The growth factor

is not part of the cell and its value is determined by the surrounding cells. The expression of CycD changes independently of the cell's content and reflects the state of the growth factor. Depending on the expression status of CycD, we obtain two realization networks for the embedded-PBN. The first realization network is determined from Table VI when the value of CycD is equal to zero. Similarly, the second realization network is determined by setting the variable of CycD to one. To completely define the embedded-PBN, the switching probability, the perturbation probability, and the probability of selecting each constituent Boolean network have to be specified. We assume that these are known. Here, we set the switching probability and the perturbation probabilities equal to 0.001, and assume that the two constituent Boolean networks are equally likely.

We also have to specify the inter-transition interval distributions between the states to fully define the cell-cycle's SM-ARN. Although to date, the lack of sufficient time-course data has prohibited the inference of any realistic asynchronous networks; the situation is expected to improve in the future with the advent of new experimental technique. Here, we simply assume that all inter-transition intervals between states are exponentially distributed. If $\tau(\boldsymbol{z}_1, \boldsymbol{z}_2)$ is the sojourn time in state $\boldsymbol{z}_1 \in \mathcal{Z}$ before transition to state $\boldsymbol{z}_2 \in \mathcal{Z}$, then we need the rate of the transition from state $\boldsymbol{z}_1$ to state $\boldsymbol{z}_2$ to specify its distribution. We use the gene-expression data to determine the probability of the transition from state $\boldsymbol{z}_1$ to state $\boldsymbol{z}_2$ in the embedded-PBN in Equation (2.18). We assume that the rate of the transition from state $\boldsymbol{z}_1$ to state $\boldsymbol{z}_2$ is assigned such that

$$\Pr\left\{\tau(\boldsymbol{z}_1, \boldsymbol{z}_2) < \min_{\substack{\boldsymbol{z}\in\mathcal{Z} \\ \boldsymbol{z}\neq\boldsymbol{z}_1}} \tau(\boldsymbol{z}_1, \boldsymbol{z})\right\} = p_{\boldsymbol{z}_1}(\boldsymbol{z}_2). \tag{3.1}$$

In other words, the probability of the first transition out of state $\boldsymbol{z}_1$ to state $\boldsymbol{z}_2$ is equal to the transition probability $p_{\boldsymbol{z}_1}(\boldsymbol{z}_2)$. The left side of Equation (3.1) can be

determined for exponentially distributed sojourn times.

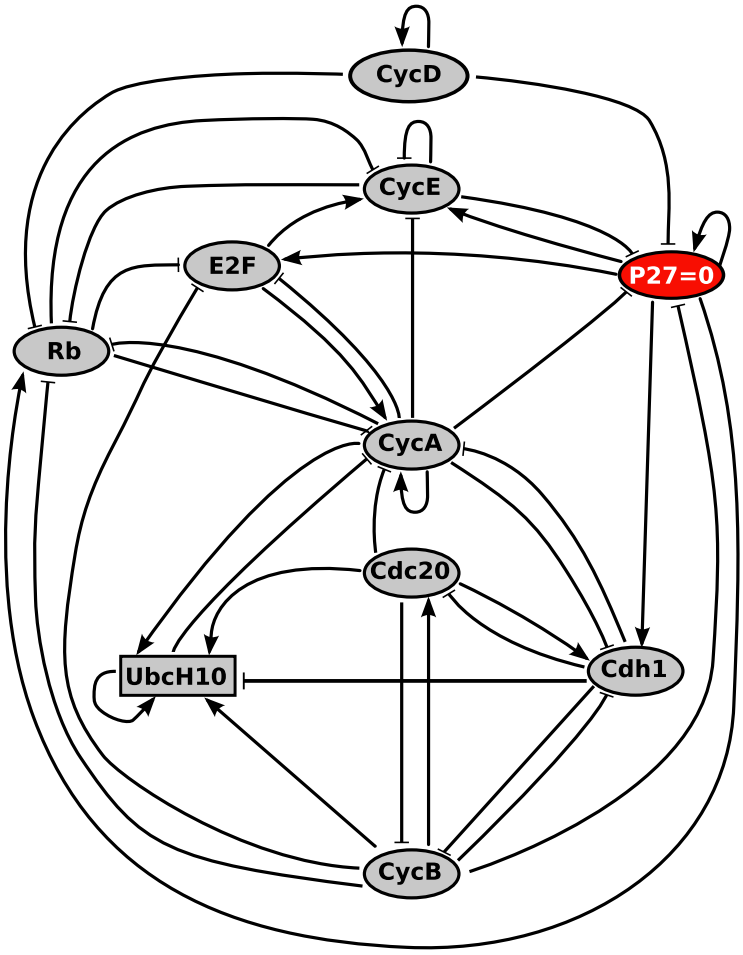
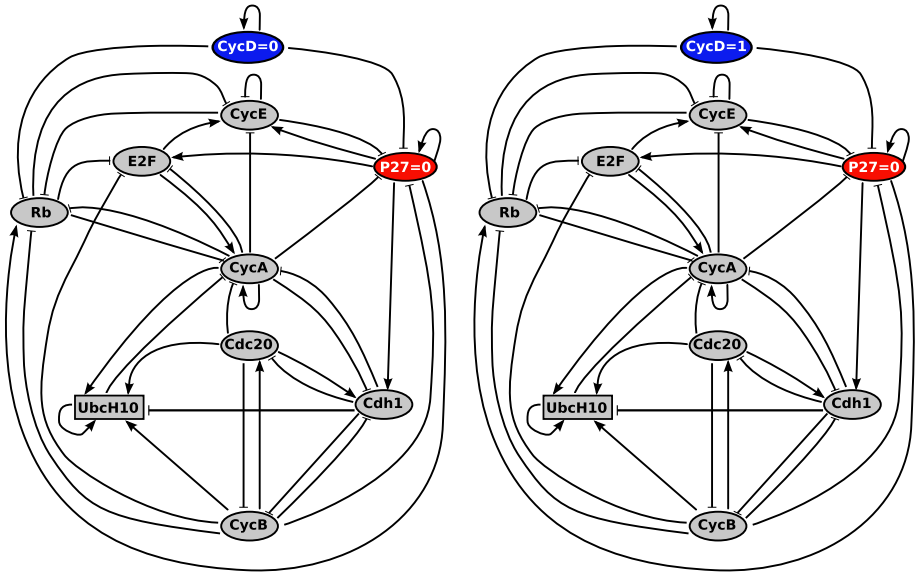


Fig. 7. The regulatory graph for the postulated mammalian cell cycle network with mutation. It is assumed that p27 is mutated and its logical state is always zero (OFF). Each node represents the activity of a key regulatory element. Blunt arrows stand for inhibitory effects, normal arrows for activations.



Constituent BN1 Constituent BN2

Fig. 8. The regulatory graphs of the two constituent Boolean networks used to construct a context-sensitive PBN for the mutated mammalian cell cycle.

CHAPTER IV

INTERVENTION IN SYNCHRONOUS MARKOVIAN REGULATORY NETWORKS *

The main objective of intervention in a regulatory network is to reduce the likelihood of encountering the undesirable states associated with aberrant cellular functions. To formulate the problem of intervention in a synchronous Markovian regulatory network, its transition probability matrix must be derived. This has been accomplished for PBNs in Chapter II. In this chapter, the task of finding the most effective intervention strategy is formulated as a classical sequential decision making problem, which is referred to as *classical intervention*.

For a pre-defined cost of intervention and a cost-per-stage function that discriminates between the states of the system, the objective of the decision maker becomes minimizing the expected total cost associated with the progression of the system. That is, given the state of the system, an effective intervention strategy identifies which action to take so as to minimize the overall cost. Consequently, the devised intervention strategy can be used as a therapeutic strategy that alters the dynamics of aberrant cells to reduce the long-run likelihood of undesirable gene-activity profiles favorable to the disease. It is evident that the intervention strategy specified by the sequential decision maker is directly affected by the form of the transition probability matrix associated with a synchronous Markovian network.

In PBNs, where genes are updated simultaneously, appropriate transition probability matrices that act on the states of the oriented graphs are sufficient to fully

describe the dynamics in equations (2.12) and (2.2). For an instantaneously random PBN, the state-space consists of gene-activity profiles while for a context-sensitive PBN, the state-space composed of ordered pairs of context and gene-activity profile.

Methods have been proposed that use information in these transition probability matrices to devise effective therapeutic strategies [6, 7]. The first proposed intervention strategies involve resetting the state of the regulatory graph of an instantaneously random PBN, as necessary, to a more desirable initial state and letting the network evolve from there [28], or changing the steady-state (long-run) probability distribution of the network by minimally altering its rule-based structure [45]. Subsequent to these early proposals, the major effort has focused on manipulating external (control) variables that affect the transition probability distributions of the regulatory graph of PBN and can, therefore, be used to desirably affect its dynamical evolution [6].

An effective intervention strategy specifies how control variables should be manipulated. The effects of an intervention strategy that is beneficial in the short-term may wear out over time. Thus, it is important to look for intervention strategies that consider the long-run effects. In the framework of PBNs, the theory of infinite-horizon Markov decision process has been employed to find optimal intervention strategies with respect to the defined objective functions [6]. Throughout this volume, we refer to this method as classical intervention.

A. Classical Intervention in Context-Sensitive Probabilistic Boolean Networks

We can formulate the task of finding the most effective intervention strategy as a sequential decision making problem, when the dynamics of a context-sensitive PBN are expressed according to (2.2). Altering the long-run likelihood of states favorable to a pathological cell functionality is the objective of the decision maker. For a pre-

defined cost of intervention and a cost-per-stage function that discriminates between the states of the system, the objective of the decision maker is to minimize the accumulated expected cost associated with the progression of the system. That is, given the state of the system, an effective intervention strategy identifies which action to take so as to minimize the overall expected cost.

To be more precise, in the presence of an external regulator, we suppose that the context-sensitive PBN has a binary intervention input $u_g(t)$ on the control gene g. The intervention input $u_g(t)$, which takes values in set $\mathcal{C} = \{0, 1\}$, specifies the action on the control gene g. Treatment alters the status of the control gene g, which can be selected from all the genes in the network. If treatment is applied, $u_g(t) = 1$, then the state of the control gene g is toggled; otherwise the state of the control gene g remains unchanged.

For the case of a single control gene g, the system evolution is represented by a stationary discrete-time equation

$$\boldsymbol{z}(t+1) = f(\boldsymbol{z}(t), u_g(t), w(t)) \qquad t = 0, 1, \ldots \tag{4.1}$$

where the state $\boldsymbol{z}(t)$ is an element of $\mathcal{Z}$; and similar to the context-sensitive PBN without control, $w(t)$ is the manifestation of uncertainties in the model.

The transition probability matrix for the controlled context-sensitive PBN can be defined easily, once the transition probability matrix of the uncontrolled system is known. Originating from a state $\boldsymbol{z}_1$, the successor state $\boldsymbol{z}_2$ is selected randomly within set $\mathcal{Z}$ according to the transition probability distribution

$$P_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u) \stackrel{\triangle}{=} \Pr(\boldsymbol{z}(t+1) = \boldsymbol{z}_2 | \boldsymbol{z}(t) = \boldsymbol{z}_1, u_g(t) = u) \tag{4.2}$$

for all $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{Z}$ and all $u \in \mathcal{C}$.

The elements of the transition probability matrix of the controlled context-

sensitive PBN given by (4.2) can then be expressed through (2.11). The value of state after intervention $\boldsymbol{z}_1 = (\kappa_1, \boldsymbol{x}_1)$ can be determined according to the status of the control signal and the value of the state prior to the intervention $\hat{\boldsymbol{z}}_1 = (\hat{\kappa}_1, \hat{\boldsymbol{x}}_1)$. Here, κ_1 is equated to $\hat{\kappa}_1$, and the value of the GAP is updated according to the value of the control signal in the devised strategy according to

$$\boldsymbol{x}_1 = (\hat{\boldsymbol{x}}_1 \oplus e_g)\mathbf{1}(u_g(\hat{\boldsymbol{z}}_1) = 1) + \hat{\boldsymbol{x}}_1\mathbf{1}(u_g(\hat{\boldsymbol{z}}_1) = 0)\cdot \tag{4.3}$$

All the 2^n elements of vector e_g are zeros except the element at the gth position, which is set to one.

To define the problem of intervention in a context-sensitive PBN, we associate a cost-per-stage $c(\boldsymbol{z}_1, \boldsymbol{z}_2, u)$ to each possible event. In general, the cost-per-stage can depend on the origin state $\boldsymbol{z}_1$, the successor state $\boldsymbol{z}_2$, and the control input u. We define the *average immediate cost* in state $\boldsymbol{z}_1$, when control u is selected, by

$$\overline{c}(\boldsymbol{z}_1, u) = \sum_{\boldsymbol{z}_2 \in \mathcal{Z}} P_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u)c(\boldsymbol{z}_1, \boldsymbol{z}_2, u). \tag{4.4}$$

We consider a discounted formulation of the expected total cost. The discounting factor, $\lambda \in (0, 1)$, ensures convergence of the expected total cost over the long-run [46]. In the case of cancer therapy, the discounting factor also emphasizes that obtaining treatment at an earlier stage is favored over later stages.

Given the system is initiated from state $\boldsymbol{z}_0$, the sequential decision maker searches for an optimal strategy π^*, one that minimizes the expected total discounted cost over the long-run progression of the PBN:

$$J_{\pi_g}(\boldsymbol{z}_0) = \lim_{N\to\infty} E\left[\sum_{t=0}^{N-1} \lambda^t c(\boldsymbol{z}(t), \boldsymbol{z}(t+1), \mu_g(\boldsymbol{z}(t), t))\middle|\boldsymbol{z}(0) = \boldsymbol{z}_0\right]\cdot \tag{4.5}$$

A strategy $\pi_g = \{\boldsymbol{\mu}_g(\cdot, 0), \boldsymbol{\mu}_g(\cdot, 1), \ldots\}$ is a sequence of decision rules $\boldsymbol{\mu}_g(\cdot, t)$ for each updating epoch t acting on control gene g, given that the initial state is $\boldsymbol{z}_0$. In

general, a decision rule $\boldsymbol{\mu}_g(\cdot, t)$ at updating epoch t selects action $u_g(t)$ according to the history of the system as well as the current state. The history $h(t)$ at the updating epoch t is composed of the sequence of previous states and actions. If the history $h(t)$ is observed at the updating epoch t, then the decision rule $\boldsymbol{\mu}_g(\cdot, t)$ determines the probability of selecting action u conditioned on the history $h(t)$ and the current state $z(t)$. We denote the set of all such strategies by Π_g, when gene g is selected as the control gene. The set $\Pi_g(M)$ is the subset of Markovian strategies within the set of all strategies Π_g defined above. A strategy is Markovian if given the current state $\boldsymbol{z}(t)$ the decision rule $\boldsymbol{\mu}_g(\cdot, t) = \boldsymbol{\mu}_g(\boldsymbol{z}(t), t) : \mathcal{Z} \rightarrow \mathcal{C}$ is independent of all the previous states and actions $h(t)$, and selects action u with probability $\boldsymbol{\mu}_g(t, u|\boldsymbol{z}(t))$ at decision epoch t. We denote the set of all stationary strategies by $\Pi_g(S)$, where a stationary strategy for control gene g is an admissible intervention strategy in $\Pi_g(M)$ of the form $\pi_g = \{\boldsymbol{\mu}_g(\cdot), \boldsymbol{\mu}_g(\cdot), \ldots\}$. Here, $\boldsymbol{\mu}_g$ denotes a time invariant decision rule. A stationary strategy is also a deterministic strategy if decision rule $\boldsymbol{\mu}_g : \mathcal{Z} \rightarrow \mathcal{C}$ is deterministic and time invariant for each updating epoch t. The set of all deterministic strategies is represented by $\Pi_g(D)$.

The sequential decision maker in the classical intervention seeks an admissible deterministic intervention strategy $\pi_g^* = \{\boldsymbol{\mu}_g^*(.), \boldsymbol{\mu}_g^*(.), \ldots\}$ in $\Pi_g(D)$ that minimizes the expected total cost $J_{\pi_g}(\boldsymbol{z}_0)$ for each initial state $\boldsymbol{z}_0 \in \mathcal{Z}$. Mathematically, an optimal classical strategy π_g^* is a solution of the optimization problem

$$\pi_g^*(\boldsymbol{z}_0) = \arg \min_{\pi_g \in \Pi_g(D)} J_{\pi_g}(\boldsymbol{z}_0), \qquad \forall \boldsymbol{z}_0 \in \mathcal{Z}. \tag{4.6}$$

An optimal strategy π_g^* is a time-invariant sequence of decision rules for each updating epoch t acting on the control gene g.

For the specifics of our formulation, an optimal strategy always exists [46]. It is

given by the fixed-point solution of the Bellman optimality equation

$$J^*(\boldsymbol{z}_1) = \min_{u \in \mathcal{C}} \left[\overline{c}(\boldsymbol{z}_1, u) + \lambda \sum_{\boldsymbol{z}_2 \in \mathcal{Z}} P_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u) J^*(\boldsymbol{z}_2) \right]. \tag{4.7}$$

Moreover, this optimal strategy is stationary and independent of the initial state $\boldsymbol{z}_0$ [46]. In general, a decision rule at updating epoch t selects action $u_g(t)$ according to the history of the system as well as the current state, but for the specifics of classical intervention an optimal strategy resides in the set of deterministic strategies $\Pi_g(D)$. Standard dynamic programming algorithms can be used to find a fixed-point of the Bellman optimality equation. In our model, gene perturbation ensures that all the states communicate with one another. Hence, the Markov decision process associated with any stationary strategy is ergodic and has a unique invariant distribution equal to its limiting distribution [34].

B. Classical Intervention in the Metastatic Melanoma Context-Sensitive Probabilistic Boolean Network

In this section, the strategy designed by classical intervention is employed to control the *Wnt*5*a*-related context-sensitive PBN in Chapter III. As noted in Section III.A, the intervention objective is to down-regulate *Wnt*5*a*. The gene *Wnt*5*a* ceasing to be down-regulated is strongly predictive of the onset of metastasis.

Here, we also compare the performance of intervention strategies designed based on the instantaneously random PBN and context-sensitive PBN when they are used to control the dynamics of context-sensitive network. If the active constituent network of the context-sensitive PBN is not observable at each instant, then one may simply use the intervention strategy design based on the equivalent instantaneously random PBN to intervene in a context-sensitive PBN. This Markov approximation to the hidden-

Markov model of context-sensitive PBN approximates the design of an intervention strategy by removing the context from the state-space of dynamic system. This means that decisions must be made without explicit knowledge of the context, and therefore without full knowledge of the system state, which is composed of a context and a gene-activity profile. As such, we would like to evaluate the effectiveness of the intervention strategy based on the instantaneously random PBN when it is used to control the actual context-sensitive PBN. This approach simplifies the task of finding intervention strategies by describing the dynamics of a context-sensitive PBN via the instantaneously random PBN with similar parameters, whose state space takes values from the set of all possible GAPs $\mathcal{X}$. We expect that the intervention devised based on instantaneously random PBN be mostly accurate when contexts switch frequently.

A number of other intervention studies based on the same data have aimed to down-regulate *Wnt5a*. This model has been used since the discovery of the relation between *Wnt5a* and metastasis. The binary nature of the up or down regulation suits our binary model. A state is desirable, i.e. belongs to $\mathcal{D}$, if $Wnt5a = 0$, and undesirable, i.e. belongs to $\mathcal{U}$, if $Wnt5a = 1$. As we mentioned earlier, application of intervention requires the designation of desirable and undesirable states, and this depends upon the existence of relevant biological knowledge. The use of *Wnt5a* is one such example where the knowledge of practitioners is incorporated in a theoretical framework. Based on our objective, the states are assigned penalties according to the cost-per-stage

$$c(\boldsymbol{z}, u) = \begin{cases} 0 & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{D}, \\ 10 & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{U}, \\ c & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{D}, \\ 10 + c & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{U}. \end{cases} \tag{4.8}$$

where c is the cost of control; and $\mathcal{U}$, $\mathcal{D}$ are the sets of undesirable and desirable states, respectively. We set $c = 1$ to make the application of intervention more plausible compared to visiting undesirable states.

Since our objective is to down-regulate $Wnt5a$, a higher penalty is assigned for states having $Wnt5a$ up-regulated. Also, for a given $Wnt5a$ status, a higher penalty is assigned when the control signal is active versus when it is not. In practice, the cost values have to mathematically capture the benefits and costs of intervention and the relative preference of states. They must be set with the help of physicians in accordance with their clinical judgement. Although this is not feasible within the realm of current medical practice, we believe that such an approach will become feasible when engineering approaches are integrated into translational medicine.

Different genes in the network (except $Wnt5a$ itself) are employed as control genes. Two intervention strategies are computed by solving the Bellman optimality equation in (4.7): (1) an optimal strategy based on the transition probability distributions of context-sensitive PBN in (2.11); (2) an approximate strategy from the equivalent instantaneously random PBN described by distributions in Equation (2.13) are derived.

The devised strategy from the transition probability distributions of context-sensitive PBN $\boldsymbol{\mu}_g^* : \mathcal{Z} \rightarrow \mathcal{C}$ specifies the action that should be taken at each time step. The second strategy, which is based on the instantaneously random PBN, only takes the GAP as its input. Since the performance of the latter strategy must be evaluated with respect to the dynamics specified by the context-sensitive PBN, we need to extend it to elements of $\mathcal{Z}$. This is achieved by simply disregarding the context element of state $\boldsymbol{z}(t)$, and determining the action based on its GAP element. We denote the resulting intervention strategy obtained through instantaneously random PBN by $\hat{\boldsymbol{\mu}}_g : \mathcal{Z} \rightarrow \mathcal{C}$.

The effectiveness of an intervention strategy can be evaluated by computing the difference between its induced cost and the cost accumulated in the absence of intervention. For the $Wnt5a$-related context-sensitive PBN, we estimate the expected total discounted cost $\bar{J}_{\boldsymbol{\mu}_g^*}$ induced by the given optimal strategy $\boldsymbol{\mu}_g^*$. To this end, we generate synthetic time-course data for a number of time steps from the transition probability matrix of the $Wnt5a$-related context-sensitive PBN, while intervening based on optimal strategy $\boldsymbol{\mu}_g^*$. We estimate the total cost by accumulating the discounted cost of each state given the action at that state. This procedure is repeated for a number of random initial states, and the average of the induced total discounted costs is computed. Following a similar procedure, the approximate strategy $\hat{\boldsymbol{\mu}}_g$ is applied to the system, and the average total discounted cost $\bar{J}_{\hat{\boldsymbol{\mu}}_g}$ is computed. Finally, we compute the average total discounted cost $\bar{J}$ for time-course data when no intervention is applied.

We consider the percentage of reduction in the average total discounted cost as a performance metric. The normalized gain obtained by each intervention strategy is taken as the immediate consequence of the intervention formulation. This metric is defined as the difference between the average total discounted cost before and after intervention, normalized by the cost before intervention. The normalized gain corresponding to the optimal strategy $\boldsymbol{\mu}_g^*$ is

$$\Delta J_g^{\mathrm{E}} = \frac{\bar{J} - \bar{J}_{\boldsymbol{\mu}_g^*}}{\bar{J}}, \tag{4.9}$$

and the normalized gain corresponding to the strategy derived from the approximate method $\hat{\boldsymbol{\mu}}$ is

$$\Delta J_g^{\mathrm{A}} = \frac{\bar{J} - \bar{J}_{\hat{\boldsymbol{\mu}}_g}}{\bar{J}}. \tag{4.10}$$

Figure 9 depicts the normalized gains when the optimal and approximate strate-

gies for each control gene are used to intervene in the context-sensitive PBN. To compute the normalized gains, we computed the costs for ten thousand trajectories of length two hundred thousand. As we expected, the optimal strategy outperforms the approximate strategy significantly for all the control genes. Moreover, for the best control gene $S100p$, the difference between the two strategies is the greatest.

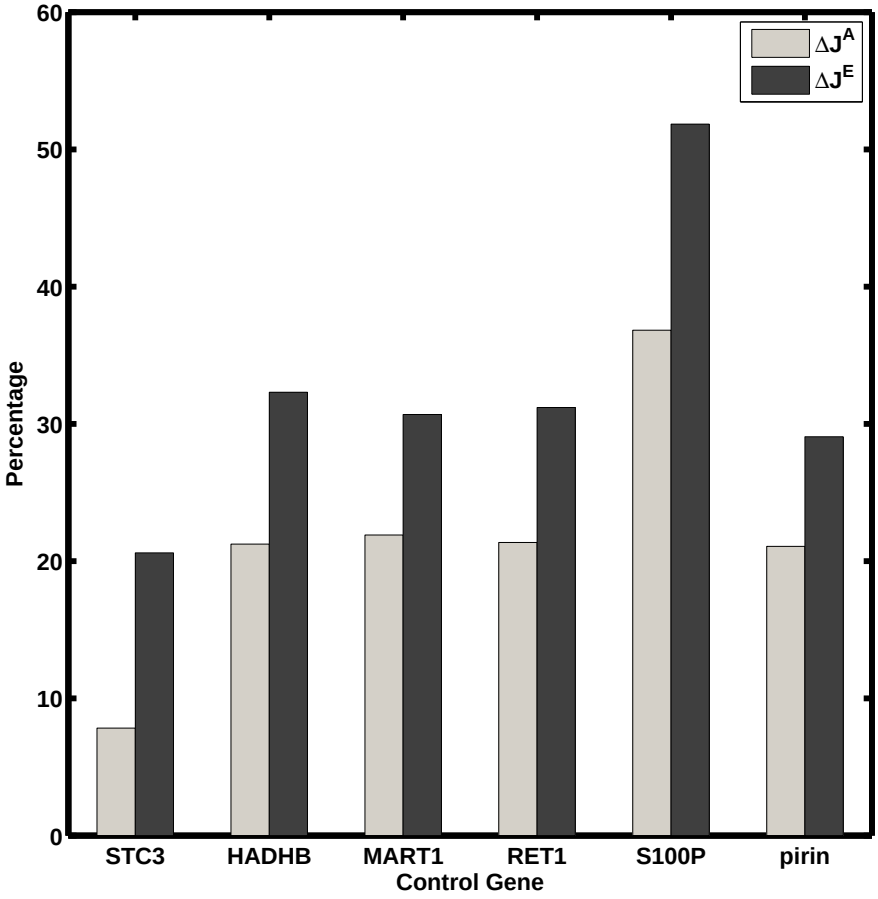


Fig. 9. ΔJ_g^E and ΔJ_g^A are computed for the Wnt5a network for various control genes.

As a byproduct of the intervention formulation, we also consider the effect of an intervention strategy $\boldsymbol{\mu}_g$ on the amount of change in the steady-state probability of undesirable states before and after the intervention. For each set of constituent

networks and for a given switching probability, we compute ΔP_g^{E} and ΔP_g^{A}. These are the normalized reduction in the total probability of visiting undesirable states in the long run for a given context-sensitive PBN when strategies $\boldsymbol{\mu}_g^*$ and $\hat{\boldsymbol{\mu}}_g$ are applied to the context-sensitive PBN, respectively. In other words, we define

$$\Delta P_g^E = \frac{\sum_{\boldsymbol{z}\in\mathcal{U}} \pi(\boldsymbol{z}) - \sum_{\boldsymbol{z}\in\mathcal{U}} \pi_{\boldsymbol{\mu}_g^*}(\boldsymbol{z})}{\sum_{\boldsymbol{z}\in\mathcal{U}} \pi(\boldsymbol{z})}, \tag{4.11}$$

and

$$\Delta P_g^A = \frac{\sum_{\boldsymbol{z}\in\mathcal{U}} \pi(\boldsymbol{z}) - \sum_{\boldsymbol{z}\in\mathcal{U}} \pi_{\hat{\boldsymbol{\mu}}_g}(\boldsymbol{z})}{\sum_{\boldsymbol{z}\in\mathcal{U}} \pi(\boldsymbol{z})}, \tag{4.12}$$

where $\pi_{\boldsymbol{\mu}_g^*}(\boldsymbol{z})$ is the probability of being in state $\boldsymbol{z}$ in the long run under optimal strategy $\boldsymbol{\mu}_g^*$; $\pi_{\hat{\boldsymbol{\mu}}_g}(\boldsymbol{z})$ is the probability of being in state $\boldsymbol{z}$ in the long run under approximate strategy $\hat{\boldsymbol{\mu}}_g$; and $\pi(\boldsymbol{z})$ is the probability of being in state $\boldsymbol{z}$ in the long run when no control is applied.

Figure 10 depicts the effects of the optimal and approximate strategies on the normalized reduction in the aggregated long-run probability of visiting undesirable states ΔP_g^{E} and ΔP_g^{A}, respectively. Here, the strategy based on the $S100p$ outperforms the strategies devised for other control genes. Note that the performance differences are not significant for most of the control genes. In particular, one should not draw any conclusions from the fact that ΔP_g^{E} is slightly less than ΔP_g^{A} in a couple of cases. The intervention strategy is designed to minimize the total cost and the improvement in the steady-state behavior is a side effect of our method.

In practice, treatment options, such as chemotherapy, have detrimental side effects. A large number of interventions can cause collateral damage that reduces a patient's quality of life. Thus, we define the quantity $\Gamma_{\boldsymbol{\mu}_g}$ as the expected number of interventions when the strategy $\boldsymbol{\mu}_g$ is applied in the long run to gauge these side effects. In particular, Γ_g^E and Γ_g^A are the expected numbers of executed interven-

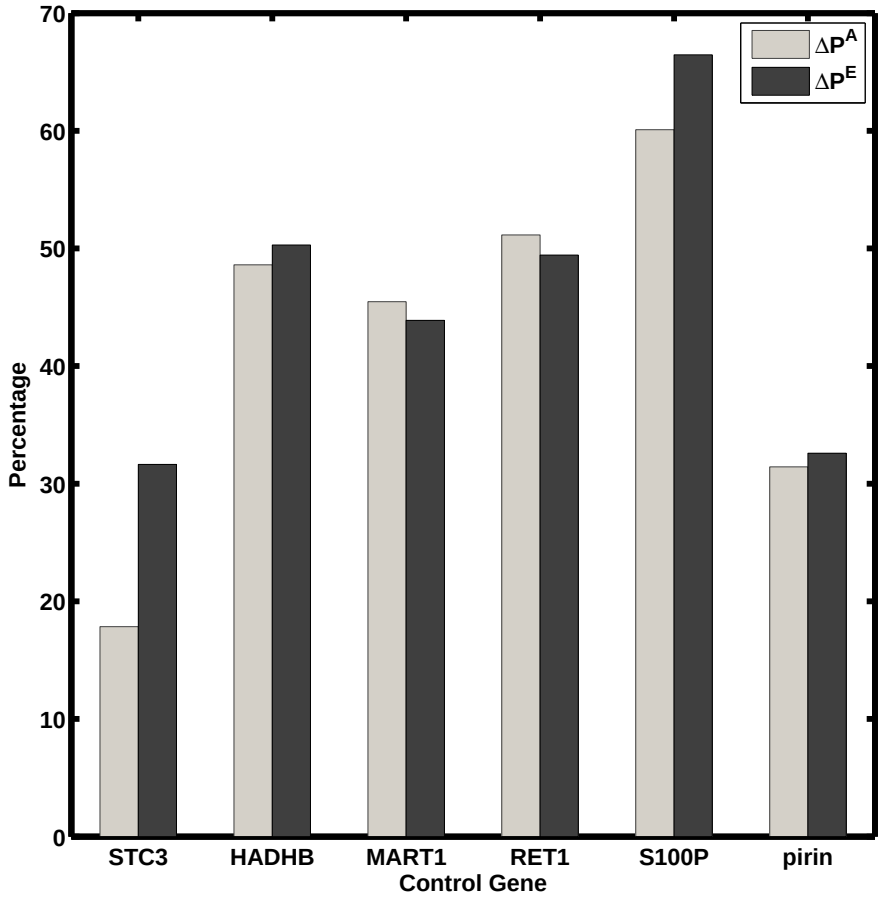


Fig. 10. ΔP_g^E and ΔP_g^A are computed for the Wnt5a network for various control genes.

tions in the long run using the optimal strategy $\boldsymbol{\mu}_g^*$ and the approximate strategy $\hat{\boldsymbol{\mu}}_g$, respectively. We define

$$\Gamma_g^E = \sum_{\boldsymbol{z}\in\mathcal{Z}} \pi_{\boldsymbol{\mu}_g^*}(\boldsymbol{z})\mathbf{1}\Big(\mu_g^*(\boldsymbol{z}) = 1\Big), \tag{4.13}$$

and

$$\Gamma_g^A = \sum_{\boldsymbol{z}\in\mathcal{Z}} \pi_{\hat{\boldsymbol{\mu}}_g}(\boldsymbol{z})\mathbf{1}\Big(\hat{\mu}_g(\boldsymbol{z}) = 1\Big), \tag{4.14}$$

where $\pi_{\boldsymbol{\mu}_g^*}(\boldsymbol{z})$ and $\pi_{\hat{\boldsymbol{\mu}}_g}(\boldsymbol{z})$ have similar definitions as in (4.11) and (4.12).

We compare the expected number of executed interventions using the difference in empirical averages, denoted by $\Delta\Gamma_g = \Gamma_g^A - \Gamma_g^E$. Figure 11 shows the difference between the expected number of executed interventions for the optimal strategy and the one derived from the approximating the network by its equivalent instantaneously random PBN. Note that the approximate strategy $\hat{\boldsymbol{\mu}}_g$ based on the most effective control gene $S100p$ applies 35% more interventions compared to the optimal strategy, while its performance is still worse.

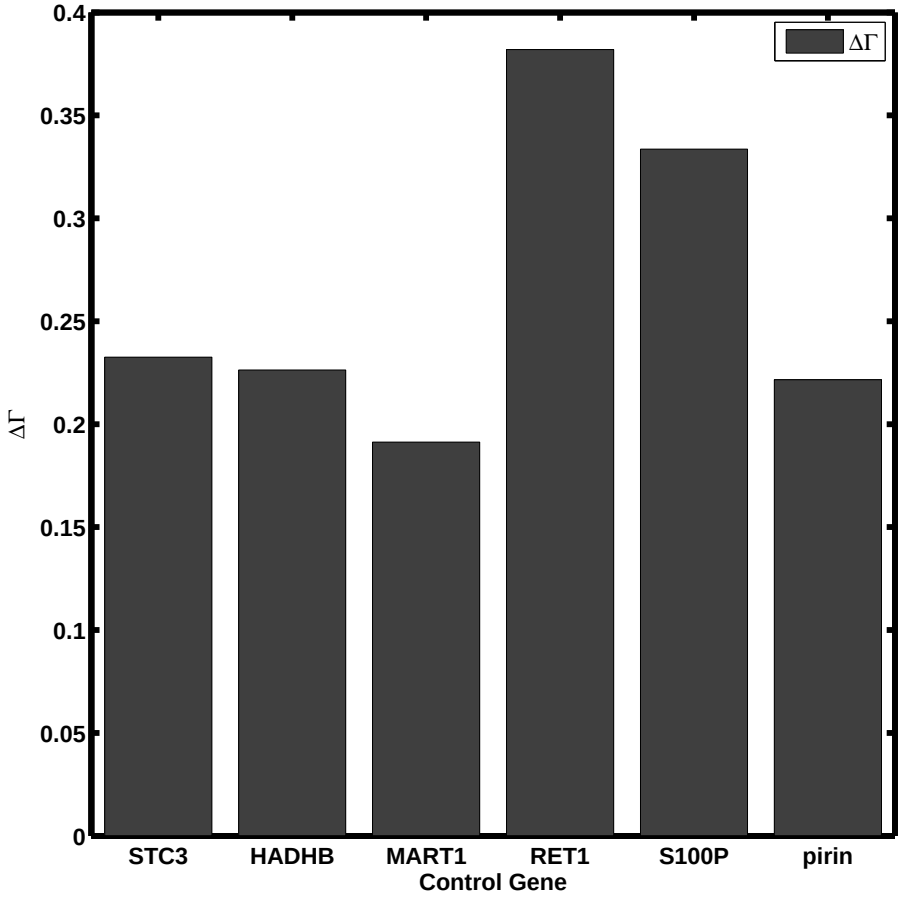


Fig. 11. $\Delta\Gamma_g$ is computed for the Wnt5a network for various control genes.

CHAPTER V

INTERVENTION IN ASYNCHRONOUS MARKOVIAN REGULATORY NETWORKS *

As it was elaborated in Chapter II, the asynchronous regulatory networks can potentially provide more effective intervention strategies. They enable us to more accurately describe the pathological cellular function of interest given our ability to perform satisfactory inference. However, we cannot readily apply classical intervention to search for effective therapeutic interventions in asynchronous networks. We introduce alternative sequential decision making techniques in this chapter. These computational tools enable us to find intervention strategies based on the asynchronous networks introduced in Chapter II.

First, we provide a methodology to derive effective intervention strategies for DA-PBNs. We show that the design of optimal intervention strategies for DA-PBNs can be mapped to the classical intervention method, although its corresponding oriented graph has a larger state space.

We resort to a synchronization method to intervene in a DA-PBN defined in Section II.B. A price in terms of computational complexity has to be paid for synchronizing the model. This synchronization procedure translates the problem of intervention in a DA-PBN to infinite-horizon discrete-time sequential decision making. This mapping augments the state-space of a PBN, specified by the logical rules of the DA-PBN, with the necessary timing history of the DA-PBN. The augmented state-

space has a considerably higher dimension. Once the oriented graph of a DA-PBN is represented by a Markov chain with augmented state-space, the classical intervention is applicable to this asynchronous model with slight modifications.

Appropriately formulating the problem of intervention in an SM-ARN, we devise an effective intervention strategy that minimizes the time that the system spends in undesirable states. Designing optimal intervention strategies based on the SM-ARN model involves results from the theory of continuous-time Markov decision processes.

We formulate the problem of intervention in SM-ARN with an arbitrary inter-transition interval distribution in Section V.B. We find optimal strategies with respect to well-defined cost functions that minimizes the duration that the system spends in undesirable states specified by rate-of-cost functions.

A. Intervention in Deterministic-Asynchronous Probabilistic Boolean Networks

Although the definition of a DA-PBN enables us to study the behavior of a regulatory network in the long run, it does not provide a systematic means for its alteration. We propose a synchronization method for DA-PBNs. The synchronization method provides a synchronous version of a DA-PBN's oriented graph. The synchronized oriented graph sets the stage for designing optimal intervention strategies to effectivey alter the dynamical behavior of DA-PBNs.

As we explained in Chapter II, to study the dynamical behavior of an instantaneously random PBN, the gene-activity profile is considered as the state of its oriented graph. The state space of a context-sensitive PBN is composed of ordered pairs of context and gene-activity profile in the set $\mathcal{Z}$.

In the synchronization of a DA-PBN, we augment the state of the oriented graph and define the *augmented logical state* $\hat{\boldsymbol{z}}(t)$. To synchronize the oriented graph of a

DA-PBN, we encode all the dynamic steps within an interval of duration equal to the least common multiple (LCM) of all the updating periods. The least common multiple of all the updating periods a_{li} in $\{\boldsymbol{\Theta}_l\}_{l=1}^{k}$, for $l \in \{1, \ldots, k\}$ and $i \in \{1, \ldots, n\}$,

$$\xi = LCM\left(a_{11}, \ldots, a_{1n}, \ldots, a_{k1}, \ldots, a_{kn}\right) \tag{5.1}$$

defines the number of a new elements m added to the vector of oriented graph state $\boldsymbol{z} \in \mathcal{Z}$. The integer m is the smallest integer larger than the logarithm to the base 2 of ξ:

$$m = \lceil \log_2(\xi) \rceil \,. \tag{5.2}$$

The value of m determined by equation (5.2) is a non-optimal number of elements required to distinguish all the time steps within one ξ. Hence, the augmented logical state of a DA-PBN at each time step t is composed of the context of the DA-PBN $\kappa(t)$, the GAP $(x_1(t), \ldots, x_n(t))$, and additional m new elements $(x_{n+1}(t), \ldots, x_{n+m}(t))$:

$$\hat{\boldsymbol{z}}(t) = (\kappa(t), x_1(t), \ldots, x_n(t), x_{n+1}(t), \ldots, x_{n+m}(t)) \,. \tag{5.3}$$

Figure 12 shows the time instants at which the genes of a hypothetical three-gene DA-PBN are updated. The updating function θ_{l1} of x_1 has the parameters $(a_{l1} = 2, b_{l1} = 1)$. Similarly, the parameters of the updating functions of genes x_2 and x_3 are $(a_{l2} = 2, b_{l2} = 0)$ and $(a_{l3} = 3, b_{l3} = 0)$, respectively. The pattern of updates is repeated after each 6 updating instants. We can use three extra elements to code all the instants in the duration of $\xi = 6$.

The evolution of a synchronized oriented graph with its augmented logical state-space can be modeled by a stationary discrete-time equation

$$\hat{\boldsymbol{z}}(t+1) = f\left(\hat{\boldsymbol{z}}(t), w(t)\right), \tag{5.4}$$

for $t = 0, 1, \ldots$, where the augmented logical state $\hat{\boldsymbol{z}}(t)$ is an element of the state-space $\hat{\mathcal{Z}} = \{(c, s) : c \in \{1, \ldots, k\}, s \in \{0, 1\}^{n+m}\}$. The disturbance $w(t)$ is the manifestation of uncertainties in the DA-PBN. It is assumed that both the gene perturbation distribution and the network switching distribution are independent and identical for all time steps t.

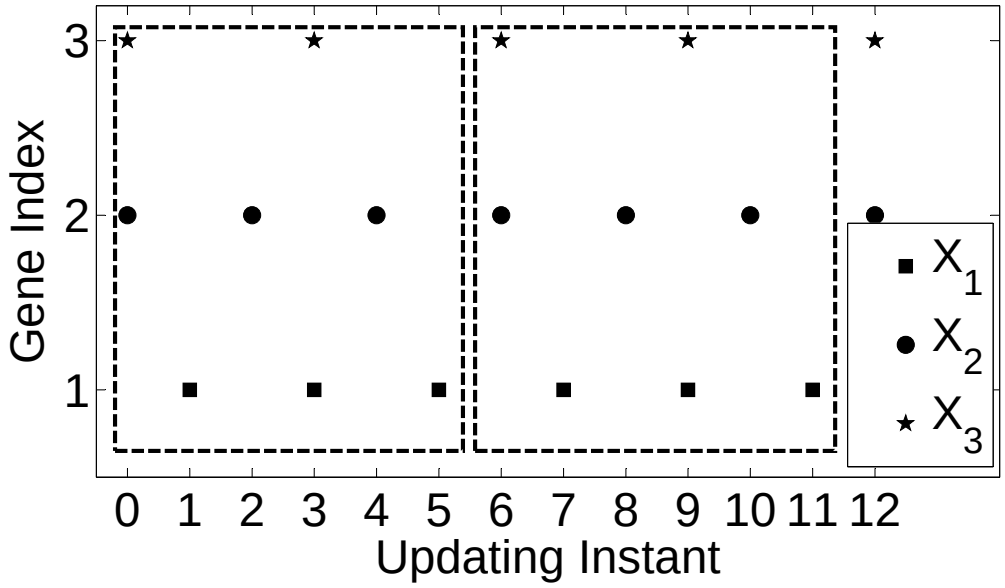


Fig. 12. Schematic of updating instants of genes of a DA-PBN with $(a_{l1} = 2, b_{l1} = 1)$, $(a_{l2} = 2, b_{l2} = 0)$ and $(a_{l3} = 3, b_{l3} = 0)$. The pattern of updates is repeated at each LCM ξ shown with a dashed-line box. Each marker indicates the instant in which the corresponding gene updates its value.

Hence, an n-gene DA-PBN is modeled as a synchronous context-sensitive PBN with the augmented state-space. The oriented graph of the synchronous context-sensitive PBN with $n+m$ elements is a Markov chain with $(2^{n+m} \times k)$ states. Hence, the oriented graph of the system described by equation (5.4) can be represented by a Markov chain [31]. Originating from an augmented logical state $\hat{\boldsymbol{z}}_1 \in \hat{\mathcal{Z}}$, the successor

augmented logical state $\hat{\boldsymbol{z}}_2 \in \hat{\mathcal{Z}}$ is selected randomly within the set $\hat{\mathcal{Z}}$ according to the transition probability

$$P_{\hat{\boldsymbol{z}}_1}(\hat{\boldsymbol{z}}_2) \stackrel{\triangle}{=} \Pr\left(\hat{\boldsymbol{z}}(t+1) = \hat{\boldsymbol{z}}_2 | \hat{\boldsymbol{z}}(t) = \hat{\boldsymbol{z}}_1\right), \tag{5.5}$$

for all $\hat{\boldsymbol{z}}_1$ and $\hat{\boldsymbol{z}}_2$ in $\hat{\mathcal{Z}}$. Gene perturbation insures that all the states in the Markov chain communicate with one another. Hence, the finite-state Markov chain is ergodic and has a unique invariant distribution equal to its limiting distribution [34].

Now that the dynamical behavior of a DA-PBN is described by a Markov chain, the theory of Markov decision processes can be utilized to find an optimal sequence of actions similar to the method developed in Section IV.A.

Reducing the likelihood of visiting undesirable augmented logical states in the long-run is the objective of the intervention problem. We suppose that the DA-PBN has an external binary control inputs u_g. The control $u_g(t)$ takes value in set $\mathcal{C} = \{0, 1\}$ at each updating instant t. In the presence of external control, the system evolution in (5.4) can be modeled by a discrete-time equation

$$\hat{\boldsymbol{z}}(t+1) = f\left(\hat{\boldsymbol{z}}(t), u(t), w(t)\right) \qquad \text{for} \quad t = 0, 1, \ldots \tag{5.6}$$

Optimal intervention in the DA-PBN is then modeled as a classical intervention with $(2^{n+m} \times k)$ states, the state $\hat{\boldsymbol{z}}(t)$ at any time step t being an augmented logical state. Originating from state $\hat{\boldsymbol{z}}_1$, the successor state $\hat{\boldsymbol{z}}_2$ is selected randomly within the set $\hat{\mathcal{Z}}$ according to the transition probability

$$P_{\hat{\boldsymbol{z}}_1}(\hat{\boldsymbol{z}}_2; u) \stackrel{\triangle}{=} \Pr\left(\hat{\boldsymbol{z}}(t+1) = \hat{\boldsymbol{z}}_2 | \hat{\boldsymbol{z}}(t) = \hat{\boldsymbol{z}}_1, u_g(t) = u\right). \tag{5.7}$$

We associate a cost-per-stage $c(\hat{\boldsymbol{z}}_1, \hat{\boldsymbol{z}}_2, u)$ to each intervention in the system. The cost-per-stage could depend on the origin state $\hat{\boldsymbol{z}}_1$, the successor state $\hat{\boldsymbol{z}}_2$, and the control input u. We also assume that the cost-per-stage is stationary and bounded

for all states $\hat{\boldsymbol{z}}_1$, $\hat{\boldsymbol{z}}_2$, and all controls $u \in \mathcal{C}$. The cost of a transition from a desirable state to an undesirable state is the highest, and the cost of a transition from an undesirable state to a desirable state is the lowest.

To define the intervention problem, we consider the discounted cost formulation, per argument in Chapter IV. An optimal intervention strategy is then determined applying the same procedure described in Section IV.A to the Markov decision process defined here.

The intervention problem in a DA-PBN has a discrete-time formulation. On the contrary, as we will show in the next section, the intervention problem in an SM-ARN has a continuous-time formulation. The objective of intervention in the discrete-time problem is to reduce the chance of visiting undesirable states. Since the time between two consecutive epochs of a DA-PBN is fixed, the effect of intervention is equivalent to the reduction of the time spent in undesirable states.

B. Intervention in Semi-Markov Asynchronous Regulatory Networks

In considering the stochastic control of an SM-ARN, we suppose that the SM-ARN has a binary control input, so $u_g(t) \in \mathcal{C} = \{0, 1\}$ describes the complete status of the control input affecting the control gene g at time t. In the presence of external control, the SM-ARN is modeled as a semi-Markov decision process.

At any time t, the state $\boldsymbol{z}(t)$ is selected from the set $\mathcal{Z}$. Originating from state $\boldsymbol{z}_1$, the successor state $\boldsymbol{z}_2$ is selected randomly within the set $\mathcal{Z}$ according to the transition probability

$$p_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u) \stackrel{\triangle}{=} \Pr\left(\boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_2 | \boldsymbol{z}(t_k) = \boldsymbol{z}_1, u_g(t_k) = u\right), \tag{5.8}$$

for all $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ in $\mathcal{Z}$ and for all u in $\mathcal{C}$. Moreover, the inter-transition interval

distribution is also a function of control $u_g(t)$:

$$P_{\boldsymbol{z}_1\boldsymbol{z}_2}(\tau; u) \triangleq \Pr\left(\tau_{k+1} \leq \tau | \boldsymbol{z}(t_k) = \boldsymbol{z}_1, \boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_2, u_g(t_k) = u\right), \tag{5.9}$$

for all states $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ in $\mathcal{Z}$, and all actions u in $\mathcal{C}$. The expressions (5.8) and (5.9) are defined based on expressions (2.18) and (2.17), respectively.

We associate a rate-of-cost $c(\boldsymbol{z}(t), u_g(t))$ for sojourning in state $\boldsymbol{z}(t)$ per unit of time while the action $u_g(t)$ is selected. Considering consecutive epoch times t_k and t_{k+1}, the rate-of-cost $c(\boldsymbol{z}(t), u_g(t))$ is constant for all $t_k \leq t < t_{k+1}$. It is equal to $c(\boldsymbol{z}, u)$, whenever $\boldsymbol{z}(t_k) = \boldsymbol{z}$ and $u_g(t_k) = u$. The rate-of-cost of undesirable states is higher than those for desirable states. We also consider the cost of applying a control action, which increases the rate-of-cost of each state.

Figure 13 shows several epoch times of a hypothetical three-gene SM-ARN. We assume that the undesirable states are the ones with an up-regulated gene in the most significant position in the GAP. We then assign lower rate-of-costs to desirable states 0 through 3 compared to the undesirable states 4 through 7. Given that r_1 and r_2 are the rate-of-costs when the model is in undesirable and desirable states, respectively, the cost $(t_2 - t_1)r_2$ gained between two epoch times t_1 and t_2 is higher than the cost $(t_6 - t_5)r_2$ gained between the two epoch times t_5 and t_6.

We desire an effective intervention policy that minimizes the accumulated cost over time. In other words, we seek an intervention strategy to reduce the time spent in undesirable states with higher rate-of-cost compared to desirable states with lower rate-of-cost. In practice, the rate-of-cost have to capture the relative preferences for the different states. For instance, the cost gained between the two epoch times t_6 and t_7 may need to be greater than the cost gained between the two epoch times t_1 and t_2. In order to penalize the sojourn time in undesirable states, the ratio of r_1 to r_2 should be large enough to capture the relative preference for the desirable states.

If the intervals between any two epoch times in Figure 13 were equal, then the problem of intervention in an SM-ARN would reduce to the intervention problem in PBNs. In this intervention problem, the objective is to reduce the number of visits to undesirable states, because the sojourn time in all states is the same, so that reducing the number of visits to undesirable states is directly equivalent to reducing the amount of time spent in these states.

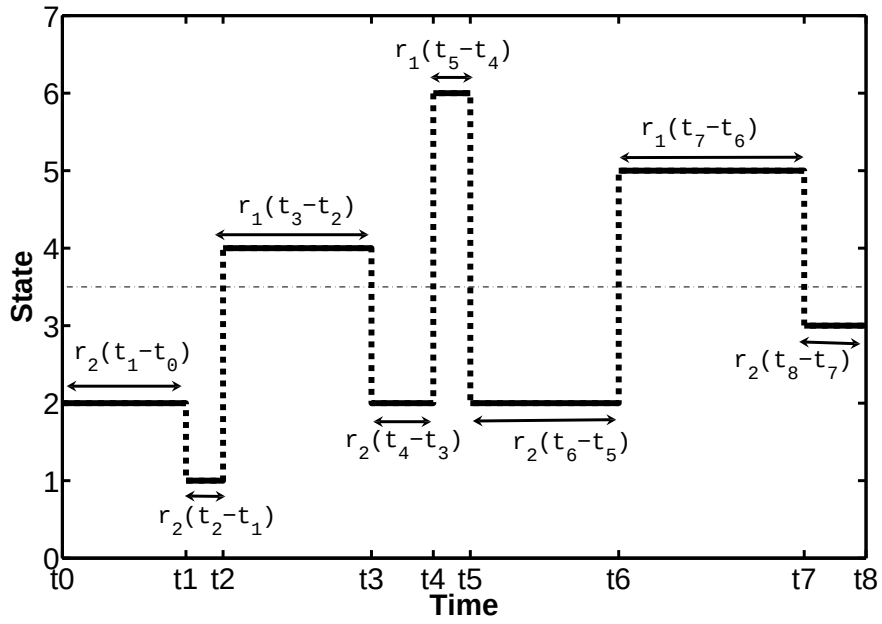


Fig. 13. Schematic of transitions in a hypothetical three-gene SM-ARN along with their epoch times and cost during each sojourn interval. The total cost between two epoch times t_1 and t_2 is less than the total cost between two epoch times t_5 and t_6.

For the reasons articulated in Chapter IV, we consider a discounted cost formulation to define the expected total cost. If $\lambda \in (0, 1)$ is the discounted factor per unit time and we divide the time unit to small intervals δ, then at each interval the discount is λ/δ, given the initial value is 1. Hence, $(1 - \frac{\lambda}{\delta})^{\delta t}$ represents the discount

over t units of time. As δ goes to zero, the discount goes to $e^{-\lambda t}$.

Among all admissible deterministic strategies $\Pi_g(D)$, the decision maker finds a strategy $\pi_g = \{\boldsymbol{\mu}_g(\cdot), \boldsymbol{\mu}_g(\cdot), \ldots\}$, where $\boldsymbol{\mu}_g : \mathcal{Z} \to \mathcal{C}$ is the decision rule at time t, that minimizes the expected total discounted cost. The expected total discounted cost, given the policy π_g and the initial state $\boldsymbol{z}_0$, is

$$J_{\pi_g}(\boldsymbol{z}_0) = \lim_{N\to\infty} E\left[\int_0^{t_N} e^{-\lambda t} c\left(\boldsymbol{z}(t), \mu_g(\boldsymbol{z}(t))\right) dt \middle| \boldsymbol{z}(t_0) = \boldsymbol{z}_0\right], \tag{5.10}$$

where t_N is the N-th epoch time. We seek a strategy π_g^* that minimizes the value function for each state $\boldsymbol{z}_0$. An optimal intervention strategy is a solution of the continuous-time decision making problem

$$\pi_g^*(\boldsymbol{z}_0) = \arg \min_{\pi_g \in \Pi_g(D)} J_{\pi_g}(\boldsymbol{z}_0), \qquad \forall \boldsymbol{z}_0 \in \mathcal{Z}. \tag{5.11}$$

Intervention using the strategy π_g^* increases the time spent in desirable states determined through appropriate assignment of rate-of-costs $c(\boldsymbol{z}(t), u(t))$ to each state-action pair $(\boldsymbol{z}(t), u(t))$. We next present the solution to the optimization problem (5.11).

Using the inter-transition interval distributions in (5.9) and the transition probability distributions in (5.8), one can define the joint transition distribution of an inter-transition interval and the successor state, given the current state and control:

$$Q_{\boldsymbol{z}_1}(\tau, \boldsymbol{z}_2; u) \stackrel{\triangle}{=} \Pr\left(\tau_{k+1} \le \tau, \boldsymbol{z}(t_{k+1}) = \boldsymbol{z}_2 | \boldsymbol{z}(t_k) = \boldsymbol{z}_1, u_g(t_k) = u\right). \tag{5.12}$$

Consequently, the expected cost of a single transition from state $z(t_k) = \boldsymbol{z}_1$ and control $u_g(t_k) = u$

$$C(\boldsymbol{z}_1, u) \stackrel{\triangle}{=} E\left[\int_0^{\tau} e^{-\lambda t} c(\boldsymbol{z}_1, u) dt\right] \tag{5.13}$$

can be computed. Here, we have

$$C(\boldsymbol{z}_1, u) = c(\boldsymbol{z}_1, u) E\left[\int_0^{\tau} e^{-\lambda t} dt\right]. \tag{5.14}$$

Furthermore, we get

$$C(\boldsymbol{z}_1, u) = c(\boldsymbol{z}_1, u) E_{\boldsymbol{z}_2}\left[E_{\tau}\left[\int_0^{\tau} e^{-\lambda t} dt\right] \middle| \boldsymbol{z}_2\right], \tag{5.15}$$

which in turn can be written as

$$C(\boldsymbol{z}_1, u) = c(\boldsymbol{z}_1, u) \sum_{\boldsymbol{z}_2 \in \mathcal{Z}} p_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u) \int_0^{\infty} \left(\int_0^{\tau} e^{-\lambda t} dt\right) \frac{dQ_{\boldsymbol{z}_1}(\tau, \boldsymbol{z}_2; u)}{p_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u)}, \tag{5.16}$$

that can be simplified as

$$C(\boldsymbol{z}_1, u) = c(\boldsymbol{z}_1, u) \sum_{\boldsymbol{z}_2 \in \mathcal{Z}} \int_0^{\infty} \frac{1 - e^{-\lambda \tau}}{\lambda} \, dQ_{\boldsymbol{z}_1}(\tau, \boldsymbol{z}_2; u). \tag{5.17}$$

A recursive relation exists between the value function $J^N_{\pi_N}$ of stage N

$$J^N_{\pi_N}(\boldsymbol{z}_0) = \sum_{k=0}^{N-1} E\left[\int_{t_k}^{t_{k+1}} e^{-\lambda t} c\left(\boldsymbol{z}_k, \mu_g(\boldsymbol{z}(t_k))\right) \, dt | \boldsymbol{z}(t_0) = \boldsymbol{z}_0\right] \tag{5.18}$$

and the value function $J^{N-1}_{\pi_{N-1}}$ of stage $(N-1)$ based on the definition (5.10), given the $(N-1)$-stage policy π_{N-1} is the subset of the N-stage policy $\pi_N = \{\boldsymbol{\mu}_g(\cdot, 0), \boldsymbol{\mu}_g(\cdot, 1), \ldots, \boldsymbol{\mu}_g(\cdot, N-1)\}$ when $\boldsymbol{\mu}_g(\cdot, 0)$ is excluded. We can express (5.18) as

$$J^N_{\pi_N}(\boldsymbol{z}_0) = C\left(\boldsymbol{z}_0, \mu_g(\boldsymbol{z}_0)\right) + E\left[e^{-\lambda \tau} J^{N-1}_{\pi_{N-1}}(\boldsymbol{z}_1) \middle| \boldsymbol{z}(t_0) = \boldsymbol{z}_0, \mu_g(\boldsymbol{z}(t_0)) = \mu_g(\boldsymbol{z}_0)\right] \tag{5.19}$$

using expression (5.17). This expression can be further modified to

$$J^N_{\pi_N}(\boldsymbol{z}_0) = C\left(\boldsymbol{z}_0, \mu_g(\boldsymbol{z}_0)\right) + E_{\boldsymbol{z}_1}\left[E_\tau\left[e^{-\lambda\tau}\middle|\boldsymbol{z}_1\right] J^{N-1}_{\pi_{N-1}}(\boldsymbol{z}_1)\middle|\boldsymbol{z}(t_0) = \boldsymbol{z}_0, \mu_g(\boldsymbol{z}(t_0)) = \mu_g(\boldsymbol{z}_0)\right], \tag{5.20}$$

given τ is the first random inter-transition interval and $\boldsymbol{z}_1$ is the successor state. Then we can rewritten (5.20) as

$$J^N_{\pi_N}(\boldsymbol{z}_0) = C\left(\boldsymbol{z}_0, \mu_g(\boldsymbol{z}_0)\right) + \sum_{\boldsymbol{z}_1 \in \mathcal{Z}} p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; \mu_g(\boldsymbol{z}_0)) \left(\int_0^\infty e^{-\lambda\tau} \frac{dQ_{\boldsymbol{z}_0}(\tau, \boldsymbol{z}_1; \mu_g(\boldsymbol{z}_0))}{p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; \mu_g(\boldsymbol{z}_0))}\right) J^{N-1}_{\pi_{N-1}}(\boldsymbol{z}_1). \tag{5.21}$$

Finally, by simplifying expression (5.21), we obtain a recursive relation between the value function $J^N_{\pi_N}$ of stage N and the value function $J^{N-1}_{\pi_{N-1}}$ of stage $(N-1)$;

$$J^N_{\pi_N}(\boldsymbol{z}_0) = C\left(\boldsymbol{z}_0, \mu_g(\boldsymbol{z}_0)\right) + \sum_{\boldsymbol{z}_1 \in \mathcal{Z}} \int_0^\infty e^{-\lambda\tau} dQ_{\boldsymbol{z}_0}(\tau, \boldsymbol{z}_1; \mu_g(\boldsymbol{z}_0)) J^{N-1}_{\pi_{N-1}}(\boldsymbol{z}_1). \tag{5.22}$$

where $\boldsymbol{z}(t_k) = \boldsymbol{z}_k$. Equation (5.22) can be rewritten as

$$J^N_{\pi_N}(\boldsymbol{z}_0) = C\left(\boldsymbol{z}_0, \mu_g(\boldsymbol{z}_0)\right) + \sum_{\boldsymbol{z}_1 \in \mathcal{Z}} m(\boldsymbol{z}_0, \boldsymbol{z}_1, \mu_g(\boldsymbol{z}_0)) J^{N-1}_{\pi_{N-1}}(\boldsymbol{z}_1), \tag{5.23}$$

where $m(\boldsymbol{z}_0, \boldsymbol{z}_1, u)$ is defined as

$$m(\boldsymbol{z}_0, \boldsymbol{z}_1, u) \triangleq \int_0^\infty e^{-\lambda\tau} dQ_{\boldsymbol{z}_0}(\tau, \boldsymbol{z}_1; u). \tag{5.24}$$

Equation (5.23) is similar to the Bellman optimality equation in dynamic programming algorithms, in which the expected immediate cost is replaced by $C(\boldsymbol{z}_0, u)$, which is the expected cost of a single transition from state $\boldsymbol{z}_0$ under control $\mu_g(\boldsymbol{z}_0, 0) = u$, and $\lambda \times p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; u)$ is replaced by $m(\boldsymbol{z}_0, \boldsymbol{z}_1, u)$. Hence, the optimal value function is

the unique fixed-point of the Bellman optimality equation

$$J^*(\boldsymbol{z}_0) = \min_{u \in \mathcal{C}} \left[C(\boldsymbol{z}_0, u) + \sum_{\boldsymbol{z}_1 \in \mathcal{Z}} m(\boldsymbol{z}_0, \boldsymbol{z}_1, u) J^*(\boldsymbol{z}_1) \right]. \tag{5.25}$$

Using (5.17) and (5.24), we can compute the expected single transition cost $C(\boldsymbol{z}_0; u)$ and $m(\boldsymbol{z}_0, \boldsymbol{z}_1, u)$, respectively, for any SM-ARN. These values define the Bellman optimality equation (5.25). Any numerical method that solves the classical intervention optimization, e.g. value iteration, can be used to find the fixed-point of (5.25), and also provides an optimal intervention strategy which is a solution to optimization (5.11). Here, we consider three hypothetical cases for the inter-transition interval distribution. For each case, we formulate the Bellman optimality equation (5.25).

1. Discrete Distribution

We postulate that the duration of the transcription of a specific gene is almost fixed, given the expression status of other genes in the network. Due to latent variables, we assume that this value is drawn from a set of possible values $\{\tau_{\boldsymbol{z}_0\boldsymbol{z}_1}(k,u)\}_{k=1,\dots,m}$ with probabilities $\{\varrho_{\boldsymbol{z}_0\boldsymbol{z}_1}(k,u)\}_{k=1,\dots,m}$. Using (5.17), we have

$$C(\boldsymbol{z}_0, u) = c(\boldsymbol{z}_0, u) \sum_{\boldsymbol{z}_1 \in \mathcal{Z}} \sum_{k=1}^{m} \frac{1 - \exp\left(-\lambda\, \tau_{\boldsymbol{z}_0\boldsymbol{z}_1}(k,u)\right)}{\lambda}\, p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; u)\, \varrho_{\boldsymbol{z}_0\boldsymbol{z}_1}(k,u). \tag{5.26}$$

The value of $C(\boldsymbol{z}_0, u)$ can easily be computed.

Using (5.24), we have

$$m(\boldsymbol{z}_0, \boldsymbol{z}_1, u) = \sum_{k=1}^{m} p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; u)\, \varrho_{\boldsymbol{z}_0\boldsymbol{z}_1}(k,u) \exp\left(-\lambda\, \tau_{\boldsymbol{z}_0\boldsymbol{z}_1}(k,u)\right), \tag{5.27}$$

so $m(\boldsymbol{z}_0, \boldsymbol{z}_1, u)$ can also be computed. Having (5.26) and (5.27), we appropriately formulate the Bellman optimality equation (5.25) for inter-transition interval with discrete distribution.

2. Uniform Distribution

We assume that, given the expression status of other genes in the network, we can measure the maximum and the minimum duration of transcription of a specific gene. We postulate that the inter-transition interval between two states can take any value within the range from the maximum value to the minimum value with an equal probability. The inter-transition interval between two states $\boldsymbol{z}_0$ and $\boldsymbol{z}_1$ has a uniform distribution in the interval $[c_{\boldsymbol{z}_0\boldsymbol{z}_1}(u), d_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)]$.

Using (5.17), we have

$$C(\boldsymbol{z}_0, u) = \frac{c(\boldsymbol{z}_0, u)}{\lambda} \sum_{\boldsymbol{z}_1 \in \mathcal{Z}} \left(1 - \frac{\exp\left(-\lambda\, c_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)\right) - \exp\left(-\lambda\, d_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)\right)}{\lambda\left(d_{\boldsymbol{z}_0\boldsymbol{z}_1}(u) \;-\; c_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)\right)}\right) p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; u). \tag{5.28}$$

Again the value of $C(\boldsymbol{z}_0, u)$ can easily be computed.

Using (5.24), we have

$$m(\boldsymbol{z}_0, \boldsymbol{z}_1, u) = \frac{\exp\left(-\lambda\, c_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)\right) - \exp\left(-\lambda\, d_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)\right)}{\lambda\left(d_{\boldsymbol{z}_0\boldsymbol{z}_1}(u) \;-\; c_{\boldsymbol{z}_0\boldsymbol{z}_1}(u)\right)}\, p_{\boldsymbol{z}_0}(\boldsymbol{z}_1; u), \tag{5.29}$$

so $m(\boldsymbol{z}_0, \boldsymbol{z}_1, u)$ can also be computed. Having (5.28) and (5.29), we again appropriately formulate the Bellman optimality equation (5.25) for inter-transition interval with uniform distribution.

3. Exponential Distribution

The amount of data observed from a biological system is usually limited. Instead of using the data to estimate an arbitrary inter-transition interval distribution, we can postulate a class of parametric distributions whose members can be defined with the first few moments, e.g. the expected value.

Here, we assume that the distribution of the inter-transition interval follows an exponential distribution. If all the inter-transition intervals of state $\boldsymbol{z}$ are ex-

ponentially distributed, then the sojourn time of state $\boldsymbol{z}$ possesses an exponential distribution:

$$P_{\boldsymbol{z}}(\tau, u) = 1 - e^{-\nu_{\boldsymbol{z}}(u)\,\tau} \qquad \tau \geq 0. \tag{5.30}$$

In (5.30), $\nu_{\boldsymbol{z}}(u)$ is the rate of transition from state $\boldsymbol{z}$ whenever the action is u. Practically, the rates $\nu_{\boldsymbol{z}}(u)$ are bounded for all states $\boldsymbol{z} \in \mathcal{Z}$, and all controls $u \in \mathcal{C}$.

Assuming the inter-transition interval is exponentially distributed, we use an alternative approach of *uniformization* to derive the Bellman optimality equation (5.25). Uniformization speeds up transitions that are slow on the average by allowing fictitious transitions from a state to itself, so sometimes after a transition the state may stay unchanged [46].

In uniformization, a uniform transition rate ν is assigned to all the states. The uniform transition rate ν is selected such that $\nu_{\boldsymbol{z}}(u) \leq \nu$ for all $\boldsymbol{z} \in \mathcal{Z}$ and $u \in \mathcal{C}$. Using the uniform transition rate ν, we define the set of new transition probabilities for each state of the uniformed decision making process by

$$\tilde{p}_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u) = \begin{cases} \frac{\nu_{\boldsymbol{z}_1}(u)}{\nu} p_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u), & \text{if } \boldsymbol{z}_1 \neq \boldsymbol{z}_2 \\ \frac{\nu_{\boldsymbol{z}_1}(u)}{\nu} p_{\boldsymbol{z}_1}(\boldsymbol{z}_2; u) + 1 - \frac{\nu_{\boldsymbol{z}_1}(u)}{\nu}, & \text{if } \boldsymbol{z}_1 = \boldsymbol{z}_2. \end{cases} \tag{5.31}$$

It can be shown that leaving state $\boldsymbol{z}$ at a rate $\nu_{\boldsymbol{z}}(u)$ in the original continuous-time decision making process is statistically identical to leaving state $\boldsymbol{z}$ at the faster rate ν, but returning back to $\boldsymbol{z}$ with the probability $(1 - \nu_{\boldsymbol{z}}(u)/\nu)$ in the uniformed decision making process with transition probability distributions defined in (5.31).

Since states of the continuous-time decision making process remain constant between two consecutive updating epochs, the expected total discounted cost in (5.10)

can be expressed as

$$J_{\pi_g}(\boldsymbol{z}_0) = \sum_{k=0}^{\infty} E\left[\int_{t_k}^{t_{k+1}} e^{-\lambda t} c(\boldsymbol{z}(t_k), \mu_g(\boldsymbol{z}(t_k)))\, dt \,\middle|\, \boldsymbol{z}(t_0) = \boldsymbol{z}_0\right]. \tag{5.32}$$

Considering the memoryless property of the exponential distribution, we can express (5.32) as

$$J_{\pi_g}(\boldsymbol{z}_0) = E\left[\sum_{k=0}^{\infty} \left(\frac{\nu}{\nu+\lambda}\right)^k \frac{c(\boldsymbol{z}(t_k), \mu_g(\boldsymbol{z}(t_k)))}{\lambda+\nu}\right]. \tag{5.33}$$

According to the latest form of the expected total discounted cost in (5.33), we can exploit the classical Markov decision process to determine the Bellman optimality transformation (5.25). The expected cost of a single transition is

$$C(\boldsymbol{z}, u) = \frac{c(\boldsymbol{z}, u)}{\lambda+\nu}, \tag{5.34}$$

and the value of parameter $m(\boldsymbol{z}_1, \boldsymbol{z}_2, u)$ is determined by

$$m(\boldsymbol{z}_1, \boldsymbol{z}_2, u) = \frac{\nu}{\lambda+\nu}\, \tilde{p}(\boldsymbol{z}_1, \boldsymbol{z}_2, u). \tag{5.35}$$

C. Intervention in the Mutated Mammalian Cell Cycle Semi-Markov Asynchronous Regulatory Network

In Section III.C, an SM-ARN is designed that models a mutated mammalian cell-cycle regulations. Our objective here is to avoid states with simultaneously down-regulated *CycD* and *Rb*. To this end, a sequential decision process based on the method presented in Section V.B is defined to determine intervention strategies for the SM-ARN of the mutated cell-cycle.

The rate of penalizing the states with down-regulated *Rb* and *CycD* is set to be higher than those for the states in which these two genes are not simultaneously

down-regulated. We postulate the following rate-of-cost function:

$$c(\boldsymbol{z}, u) = \begin{cases} 0, & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{D} \\ 5, & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{U} \\ 1, & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{D} \\ 6, & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{U}, \end{cases} \tag{5.36}$$

where $\mathcal{U}$ and $\mathcal{D}$ are the sets of undesirable and desirable states, respectively. A state $\boldsymbol{z}$ is desirable, i.e. belong to $\mathcal{D}$, if $(CycD,Rb) \neq (0,0)$; and undesirable, i.e. belong to $\mathcal{U}$, if $(CycD,Rb) = (0,0)$. We select an arbitrary rate-of-cost; however, the state cost and the control cost are selected so that applying the control to prevent the undesirable states is preferable in comparison to not applying control and remaining in an undesirable state. In practice, the cost values have to capture the benefits and side effects of the intervention and the relative preference of the states.

Figure 14 depicts the fraction of time that the SM-ARN spends in each state when there is no intervention. According to this figure, the aggregated fraction of time that the mutated cell-cycle SM-ARN spends in the states with simultaneously down-regulated $CycD$ and Rb is 49%.

We define Δp_g to be the percentage of the change in the fraction of time that the SM-ARN spends in the states with simultaneously down-regulated $CycD$ and Rb before and after the intervention with control gene g. In other words, we have

$$\Delta p_g = \frac{\sum_{\boldsymbol{z} \in \mathcal{U}} p(\boldsymbol{z}) - \sum_{\boldsymbol{z} \in \mathcal{U}} p_{\boldsymbol{\mu}_g^*}(\boldsymbol{z})}{\sum_{\boldsymbol{z} \in \mathcal{U}} p(\boldsymbol{z})}, \tag{5.37}$$

where similar to (2.19) $p(\boldsymbol{z})$ is the fraction of time that the SM-ARN spends in state $\boldsymbol{z}$ in the long run when no control is applied; and $p_{\boldsymbol{\mu}_g^*}(\boldsymbol{z})$ is the fraction of time being in state $\boldsymbol{z}$ in the long run under intervention with an optimal strategy $\boldsymbol{\mu}_g^*$. As a

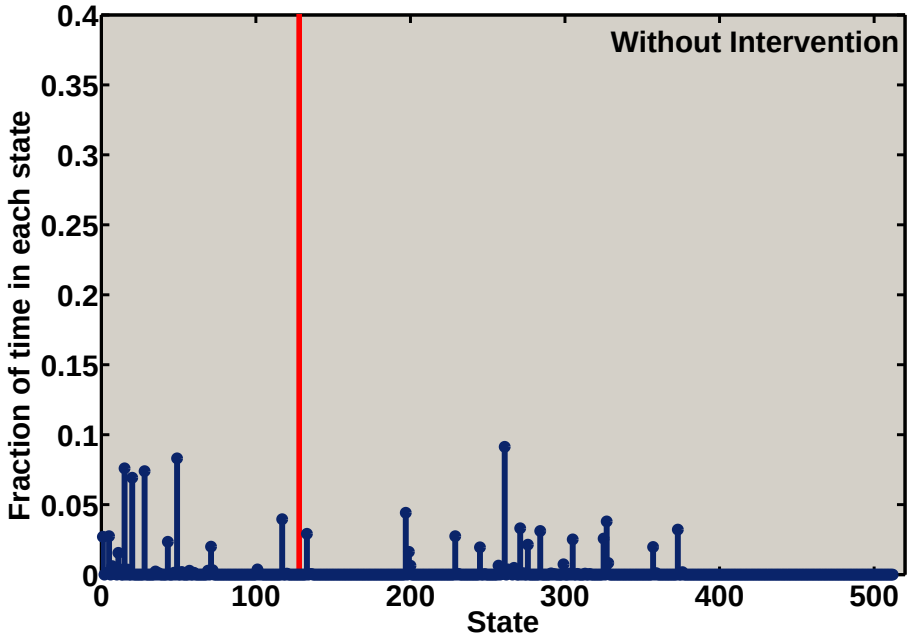


Fig. 14. The fraction of time that the SM-ARN of mammalian cell cycle spends in each state prior to intervention. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable states in $\mathcal{D}$.

performance measure, Δp_g indicates the percentage of the reduction in the fraction of time that the model spends in undesirable states in the long run.

If we assume that we can target any gene in the network as a therapeutic method, then it is natural to ask which gene should be used as a control gene to alter the behavior of the model. To this end, we find an intervention strategy for each of the genes in the network using the intervention method explained in Section V.B. Table VII lists the value of Δp_g corresponding to each control gene in the network. Among all the genes, *Rb* and *E2F* have the best performance.

Table VII. The Δp_g for the intervention strategy based on various control genes.

Control Gene g	*Rb*	*E2F*	*CycE*	*CycA*	*Cdc*20	*Cdh*1	*UbcH*10	*CycB*
Δp_g	94.2%	89.1%	71.1%	62.1%	63.5%	68.4%	59.7%	75.2%

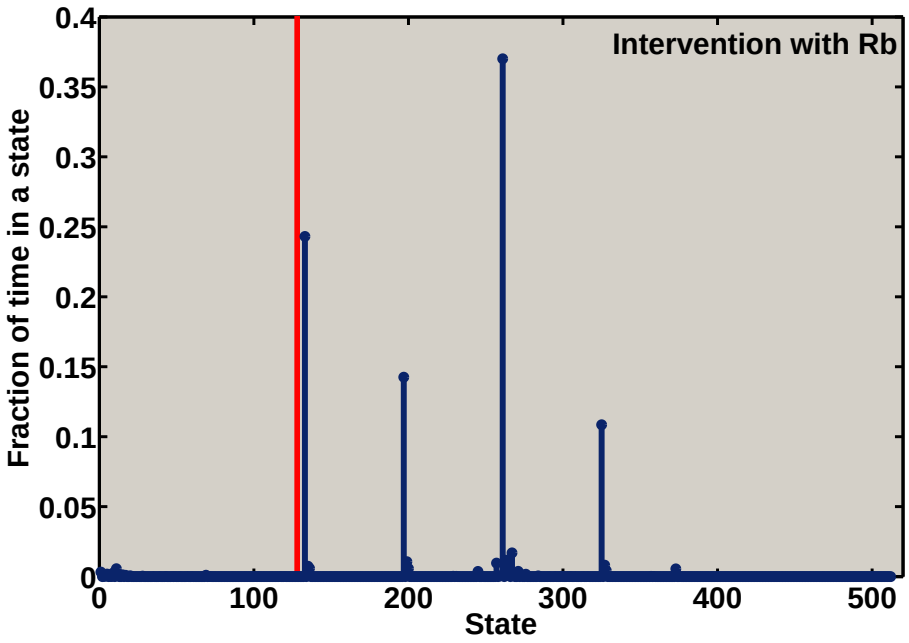


Fig. 15. The fraction of time that the SM-ARN of mammalian cell cycle spends in states after intervention using Rb as the control gene. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable states in $\mathcal{D}$.

The fraction of time that the SM-ARN of mutated mammalian cell cycle spends in states after *Rb*-based intervention is shown in Figure 15. It is clear that after intervention using *Rb* as the control gene, the fraction of time that the model spends in the undesirable states is significantly reduced. Directly using *Rb* as the control gene, the fraction of time that the model spends in the undesirable states is reduced to less than 2%.

If direct intervention based on *Rb* is not feasible, then one can use *E2F* as the control gene. According to Fig.16, in this case the system spends slightly more time in the undesirable states, but even this value is still less than 4.5%. Practically, the difference between the performances of these two control genes is insignificant. Figures 14 and 15 lead us to the conclusion that the intervention method proposed

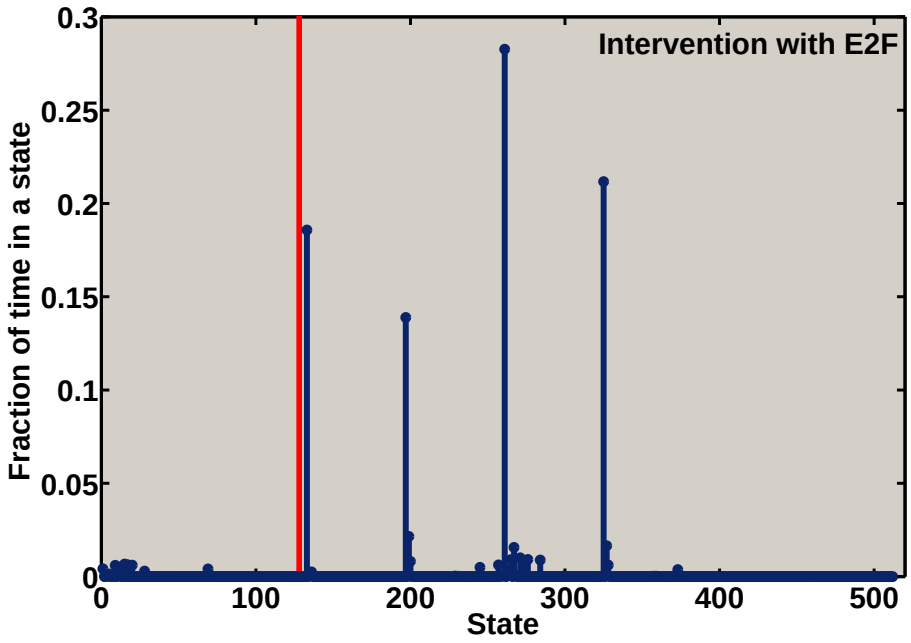


Fig. 16. The fraction of time that the SM-ARN of mammalian cell cycle spends in states after intervention using E2F as the control gene. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable states in $\mathcal{D}$.

in Section VB effectively alters the dynamics of the mutated cell-cycle SM-ARN.

CHAPTER VI

INTERVENTIONS WITH LIMITED SIDE-EFFECTS *

In classical intervention, at each epoch a devised strategy dictates whether to intervene or not in order to reduce the likelihood of undesirable states without imposing any restrictions on the quantity of applied treatments. In medical practice, however, dose intensity in a treatment is limited to mitigate the detrimental side-effects of therapy.

Here, we describe an approach that amends the unrestricted classical intervention strategy with the goal of accommodating such constraints. In the proposed approach, referred to as *constrained intervention*, the side-effects are controlled by bounding the quantity of the prescribed interventions [21].

To determine the best integrated effect consistent with a reasonable quality of life, we seek an effective therapeutic method that reduces the likelihood of states related to an undesirable cell functionality by minimizing the associated cost function, while providing an upper bound on the expected number of interventions received by a patient. As an example, we explain how to design an effective constrained intervention strategies for the mutated cell-cycle network in Section III.B.

The complete treatment of the constrained intervention is presented in the next section. To this end, we define the total constraining cost given a strategy and an initial state. Having this new constraining cost function, we reformulate the unconstrained classical intervention problem formulated in optimization (4.6) as a constrained intervention. It is shown that the objective cost function and the con-

straining cost function can be presented as a linear combination of the *occupation measure* and the cost-per-stage functions. The occupation measure can be interpreted as the probability of occupying state-action pairs in the long run, given the initial state of the PBN and the intervention strategy. Consequently, the expected number of interventions in the long run can be constrained by an upper-bound if the constrained cost-per-stage for each state-action pair is assigned as follows: $r(\boldsymbol{z}, u) = 0$ if no intervention is applied and $r(\boldsymbol{z}, u) = 1$ otherwise. This enables us to redefine the sequential decision making problem of constrained intervention as a linear program. An optimal constrained intervention strategy can then be found based on a solution of this linear program.

A. Constrained Intervention in Context-Sensitive Probabilistic Boolean Networks

Cancer treatment may include the use of chemotherapy, radiation therapy, targeted gene therapy, etc. All of these treatment options are directed at killing or eradicating cancerous cells. Unfortunately, cancer treatments may also damage healthy cells. This results in complications and harmful side-effects. It is therefore desirable to restrain the side-effects of a treatment. This goal can be achieved by enforcing an upper bound on the expected number of treatments a patient may receive during therapy.

A classical intervention strategy, devised by solving the unconstrained optimization problem (4.6), reduces the chances of visiting undesirable states; however, it does not provide a mechanism for constraining the frequency of applying treatments within the resulting intervention strategy. To address this shortcoming, constrained intervention is introduced by imposing an appropriate constraint on the optimization problem (4.6).

This is accomplished by associating constrained cost-per-stage $r(\boldsymbol{z}, u)$ with each state-action pair $(\boldsymbol{z}, u) \in \mathcal{Z} \times \mathcal{C}$. This new cost-per-stage should be defined to appropriately reflect the intended constraint. Specifically, we bound the expected number of interventions in the long run to limit the dose intensity of an intervention within a prescribed treatment.

Using constrained intervention methods, we seek an effective regulatory treatment that reduces the likelihood of visiting undesirable states in the long run, while providing an upper bound on the expected number of interventions a patient can receive. Instead of introducing a single cost function whose minimization reduces the likelihood of entering undesirable states, we consider a situation where one type of cost is minimized while keeping the other cost function below a given threshold. Posed this way, the intervention problem can be viewed as a constrained Markov decision process.

So far, a regulatory network has been modeled as a dynamic system in which decisions regarding treatment are taken sequentially. In this section, we wish to design an intervention strategy that selects treatments (actions) as a function of time and available information. For a given intervention strategy, the choice of treatments at different decision epochs may depend on the whole observed history. The choice of an intervention strategy will determine the evolution of the state of an intervened biological system in some probabilistic sense. The trajectories of the states together with the choice of treatments determine the expected cost in conjunction with the expected constrained cost that we encounter. Hence, the proposed method enables us to design therapeutic intervention strategies by defining problem dependent constraints. Although various forms of constraints are plausible, hereafter, we focus on the expected number of treatments.

The normalized expected total discounted cost, given strategy π_g, initial state

$\boldsymbol{z}_0$, and control gene g, is denoted by

$$\boldsymbol{J}_{\pi_g}(\boldsymbol{z}_0) = (1-\lambda) \times \lim_{N\to\infty} E\left[\sum_{t=0}^{N-1} \lambda^t\, c\left(\boldsymbol{z}(t), \boldsymbol{z}(t+1), \mu_g\left(\boldsymbol{z}(t), t\right)\right) \,\middle|\, \boldsymbol{z}(0) = \boldsymbol{z}_0\right]. \quad (6.1)$$

In the previous chapters, the discounted cost was defined without the normalizing constant $(1-\lambda)$. This constant does not change the method and the solution of the intervention strategy. However, using the normalizing constant has several advantages. First, it prevents the total cost from growing excessively for values of λ close to one. Second, the use of the normalization constant provides an interesting interpretation for the total cost in the constrained intervention design. This will become clear later in this section.

The vector of normalized expected total discounted costs $\boldsymbol{J}_{\pi_g} \in \mathbb{R}^{|\mathcal{Z}|}$ is called the value function. In the classical intervention problem, we seek an admissible intervention strategy π_g^* that minimizes the value function for each initial state $\boldsymbol{z}_0$, i.e.,

$$\pi_g^*(\boldsymbol{z}_0) = \arg \min_{\pi_g \in \Pi_g} \boldsymbol{J}_{\pi_g}(\boldsymbol{z}_0) \qquad \forall\, \boldsymbol{z}_0 \in \mathcal{Z}. \quad (6.2)$$

A deterministic intervention strategy devised by solving the unconstrained optimization (6.2) reduces the chance of visiting undesirable states; however, this intervention strategy does not provide a way to constrain the frequency of applying treatments within a prescribed intervention strategy. To address this shortcoming, we impose an appropriate constraint on the optimization problem (6.2) by introducing constrained intervention in Markovian regulatory networks.

For the same reasons articulated in Chapter IV, we consider a discounted formulation to define both the objective cost function and constrained cost function. To restrict the frequency of applying intervention, we associate a constrained cost-per-stage $r(\boldsymbol{z}, u)$ to each state-action pair $(\boldsymbol{z}, u)$ in the constrained formulation. The set

of all possible state-action pairs is denoted by $\mathcal{K} = \{(\boldsymbol{z}, u) : \boldsymbol{z} \in \mathcal{Z}, u \in \mathcal{C}\}$.

A constrained cost-per-stage should be defined to appropriately reflect the constraint. Here, we bound the discounted expected number of interventions in the long run. Accordingly, the normalized expected total discounted cost of the constraint, given strategy π_g, initial state $\boldsymbol{z}_0$, and control gene g is denoted by

$$\boldsymbol{H}_{\pi_g}(\boldsymbol{z}_0) = (1-\lambda) \times \lim_{N\to\infty} E\left\{ \sum_{t=0}^{N-1} \lambda^t r\left(\boldsymbol{z}(t), \mu_g\left(\boldsymbol{z}(t), t\right)\right) \,\middle|\, \boldsymbol{z}(0) = \boldsymbol{z}_0 \right\}. \tag{6.3}$$

Having the constrained cost function defined this way and the objective cost function as in (6.1), we can state the constrained intervention problem as

$$\min_{\pi_g \in \Pi_g} \boldsymbol{J}_{\pi_g}(\boldsymbol{z}_0) \qquad \text{such that } \boldsymbol{H}_{\pi_g}(\boldsymbol{z}_0) \leq C_{\text{total}}, \tag{6.4}$$

where C_{total} is the upper bound on the discounted expected number of interventions in the long run, and $\boldsymbol{z}_0$ is the initial state.

We wish to find an optimal intervention strategy π_g^* within the set of admissible strategies Π_g (not just Markovian strategies) that minimizes the value function while satisfying the constraint imposed on the discounted expected total cost. Interventions using strategy π_g^* increase the time spent in desirable states, while limiting the discounted expected number of treatments. The intervention strategy is determined through the appropriate assignments of objective cost-per-stage and constrained cost-per-stage to each state-action pair in the set $\mathcal{K}$.

Given an arbitrary strategy π_g and starting from initial state $\boldsymbol{z}_0$, the state trajectories and selected actions over time are probabilistic. Our objective is to find the expectation of the number of times that state-action pairs $(\boldsymbol{z}, u) \in \mathcal{K}$ with active intervention decision, $u = 1$, occur over the progression of the regulatory network. This value corresponds to the expected number of treatments in an intervention strategy. To this end, we denote the probability that a state-action pair $(\boldsymbol{z}, u)$ in the set of all

possible state-action pairs $\mathcal{K}$ occurs at updating epoch t as

$$\Upsilon_{\pi_g}\left(\boldsymbol{z}(t)=\boldsymbol{z},\mu_g\left(\boldsymbol{z}(t),t\right)=u\,\middle|\,\boldsymbol{z}(0)=\boldsymbol{z}_0\right). \tag{6.5}$$

We further define the normalized discounted total expected time spent in the state-action pair $(\boldsymbol{z},u)$ in the long run as

$$f_\lambda(\boldsymbol{z},u;\boldsymbol{z}_0,\pi_g)\stackrel{\triangle}{=}(1-\lambda)\times\lim_{N\to\infty}\sum_{t=0}^{N-1}\lambda^t\Upsilon_{\pi_g}\left(\boldsymbol{z}(t)=\boldsymbol{z},\mu_g\left(\boldsymbol{z}(t),t\right)=u\,\middle|\,\boldsymbol{z}(0)=\boldsymbol{z}_0\right) \tag{6.6}$$

for all $(\boldsymbol{z},u)\in\mathcal{K}$, where $\boldsymbol{z}_0$ is an initial state and π_g is a strategy in Π_g. The set

$$\mathbf{f}_\lambda(\boldsymbol{z}_0,\pi_g)=\left\{f_\lambda(\boldsymbol{z},u;\boldsymbol{z}_0,\pi_g)\,\middle|\,(\boldsymbol{z},u)\in\mathcal{K}\right\} \tag{6.7}$$

denotes a probability measure over the set of state-action pairs $\mathcal{K}$. The numbers of states and actions of regulatory networks are finite and the discounting factor λ guarantees uniform convergence of expression (6.6). The set $\mathbf{f}_\lambda(\boldsymbol{z}_0,\pi_g)$ for any initial state $\boldsymbol{z}_0$ and strategy π_g is called an *occupation measure* [47]. The occupation measure can be interpreted as the probability of occupying state-action pairs $(\boldsymbol{z},u)$ in the long run, given that the regulatory network is initially in state $\boldsymbol{z}_0$ and strategy π_g is used throughout.

The normalized discounted objective cost function (6.1) can be expressed as the expectation of the average immediate cost $\overline{c}(\boldsymbol{z},u)$, defined in (4.4), over the probability distribution defined in (6.5).

$$\begin{aligned}\boldsymbol{J}_{\pi_g}(\boldsymbol{z}_0)=(1-\lambda)\times\lim_{N\to\infty}\sum_{t=0}^{N-1}\Bigg\{&\lambda^t\sum_{(\boldsymbol{z},u)\,\in\,\mathcal{K}}\overline{c}\left(\boldsymbol{z}(t)=\boldsymbol{z},\mu_g\left(\boldsymbol{z}(t),t\right)=u\right)\\&\times\Upsilon_{\pi_g}\left(\boldsymbol{z}(t)=\boldsymbol{z},\mu_g\left(\boldsymbol{z}(t),t\right)=u\,\middle|\,\boldsymbol{z}(0)=\boldsymbol{z}_0\right)\Bigg\}\end{aligned} \tag{6.8}$$

The normalized discounted objective cost function in (6.8) can be equivalently ex-

pressed as

$$\boldsymbol{J}_{\pi_g}(\boldsymbol{z}_0) = \sum_{(\boldsymbol{z},u)\,\in\,\mathcal{K}} \Bigg\{ (1-\lambda)\times \\ \lim_{N\to\infty} \sum_{t=0}^{N-1} \Big[\lambda^t\, \overline{c}(\boldsymbol{z}(t)=\boldsymbol{z}, \mu_g\left(\boldsymbol{z}(t),t\right)=u)\times \\ \Upsilon_{\pi_g}\left(\boldsymbol{z}(t)=\boldsymbol{z}, \mu_g\left(\boldsymbol{z}(t),t\right)=u \,\middle|\, \boldsymbol{z}(0)=\boldsymbol{z}_0\right) \Big] \Bigg\}. \tag{6.9}$$

Using definition (6.6) and probability measure (6.7), we can express the latest form of the normalized discounted objective cost in Eq. (6.9) as the expectation of the average immediate objective cost with respect to the occupation measure,

$$\boldsymbol{J}_{\pi_g}(\boldsymbol{z}_0) = \sum_{(\boldsymbol{z},u)\,\in\,\mathcal{K}} f_\lambda(\boldsymbol{z},u;\boldsymbol{z}_0,\pi_g)\,\overline{c}(\boldsymbol{z},u)\cdot \tag{6.10}$$

Similarly, we can express the normalized discounted constrained cost corresponding to strategy π_g as the expectation of the constrained cost-per-stage with respect to the occupation measure

$$\boldsymbol{H}_{\pi_g}(\boldsymbol{z}_0) = \sum_{(\boldsymbol{z},u)\,\in\,\mathcal{K}} f_\lambda(\boldsymbol{z},u;\boldsymbol{z}_0,\pi_g)\,r(\boldsymbol{z},u)\cdot \tag{6.11}$$

Using (6.10) and (6.11), we can rewrite the constrained optimization problem (6.4) as

$$\begin{aligned} &\min_{\pi_g\in\,\Pi_g} \sum_{(\boldsymbol{z},u)\,\in\,\mathcal{K}} f_\lambda(\boldsymbol{z},u;\boldsymbol{z}_0,\pi_g)\,\overline{c}(\boldsymbol{z},u), \\ &\text{such that} \sum_{(\boldsymbol{z},u)\,\in\,\mathcal{K}} f_\lambda(\boldsymbol{z},u;\boldsymbol{z}_0,\pi_g)\,r(\boldsymbol{z},u) \le C_{\text{total}}. \end{aligned} \tag{6.12}$$

It is evident that the constraint in (6.12) prevents the discounted expected number of interventions in the long run from exceeding the upper-bound C_{total} if we assign

the constrained cost-per-stage for each state-action pair in $\mathcal{K}$ as

$$r(\boldsymbol{z}, u) = \begin{cases} 0, & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{Z} \\ 1, & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{Z}. \end{cases} \tag{6.13}$$

In other words, using the definition of constrained cost-per-stage in (6.13), the left side of the inequality constraint in (6.12) corresponds to the total discounted expected number of times that state-action pairs with active treatment, $u = 1$, occur under control strategy π_g. Equivalently, we can interpret this as the discounted frequency of applying treatments given a therapeutic strategy.

Several solutions for the constrained optimization problem of (6.4) are presented in [48]. We next briefly present a method to solve this constrained Markov decision process using the equivalent problem formulation of (6.12). In [48], it is shown that the set of stationary strategies $\Pi_g(S)$ is complete. In other words, if

$$L_U = \Big\{ \mathbf{f}_\lambda(\boldsymbol{z}_0, \pi_g) \,|\, \pi_g \in \Pi_g \Big\} \tag{6.14}$$

denotes the set of all the occupation measures and

$$L_{U(S)} = \Big\{ \mathbf{f}_\lambda(\boldsymbol{z}_0, \pi_g) \,|\, \pi_g \in \Pi_g(S) \Big\} \tag{6.15}$$

denotes the set of occupation measures generated by stationary strategies only, then $L_U = L_{U(S)}$. Further, let $\Xi_\lambda(\boldsymbol{z}_0)$ be defined as the set of $\mathbb{R}^{|\mathcal{K}|}$ vectors
$\boldsymbol{\alpha} = (\alpha(0,0), \alpha(0,1), \cdots, \alpha(k \times 2^n - 1, 0), \alpha(k \times 2^n - 1, 1))$ that satisfy

$$\begin{aligned} &\sum_{(\boldsymbol{z},u) \in \mathcal{K}} \alpha(\boldsymbol{z}, u) \Big(\mathbf{1}(\boldsymbol{z} = \boldsymbol{z}') - \lambda \, P_{\boldsymbol{z}}(\boldsymbol{z}'; u) \Big) = (1 - \lambda) \, \mathbf{1}(\boldsymbol{z}_0 = \boldsymbol{z}') \quad \text{for all } \boldsymbol{z}' \in \mathcal{Z}, \\ &\alpha(\boldsymbol{z}, u) \geq 0 \qquad \forall (\boldsymbol{z}, u) \in \mathcal{K}, \end{aligned} \tag{6.16}$$

where $\mathbf{1}(\cdot)$ is indicator function. If $\boldsymbol{\alpha} \in \Xi_\lambda(\boldsymbol{z}_0)$, then one can verify that $\sum_{(\boldsymbol{z},u) \in \mathcal{K}} \alpha(\boldsymbol{z}, u) = 1$ by summing the first constraint on $\boldsymbol{\alpha}$ in the definition of $\Xi_\lambda(\boldsymbol{z}_0)$ over all $\boldsymbol{z}' \in \mathcal{Z}$.

Hence, the elements of any $\boldsymbol{\alpha}$ satisfying the constraints in (6.16) constitute a probability measure on $\mathcal{K}$.

It has been shown that $L_{U(S)} = \overline{L}_{U(D)}$, where $L_{U(D)} = \left\{ \mathbf{f}_\lambda(\boldsymbol{z}_0, \pi_g) | \pi_g \in \Pi_g(D) \right\}$ and $\overline{L}_{U(D)}$ is the closed convex hull of deterministic strategies [48]. Moreover, the closed convex hull of deterministic strategies $\overline{L}_{U(D)}$ is equal to the closed polytope specified by $\Xi_\lambda(\boldsymbol{z}_0)$. Hence, from the definition in (6.16) and the constrained cost formulation (6.11), we can find an optimal strategy that satisfies (6.12) by solving the following linear program:

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^{|\mathcal{K}|}} \sum_{(\boldsymbol{z},u) \in \mathcal{K}} \alpha(\boldsymbol{z},u)\, \overline{c}(\boldsymbol{z},u),$$

such that

$$\sum_{(\boldsymbol{z},u) \in \mathcal{K}} \alpha(\boldsymbol{z},u) \Big(\mathbf{1}(\boldsymbol{z} = \boldsymbol{z}') - \lambda\, P_{\boldsymbol{z}}(\boldsymbol{z}';u) \Big) = (1-\lambda)\, \mathbf{1}(\boldsymbol{z}_0 = \boldsymbol{z}') \quad \text{for all } \boldsymbol{z}' \in \mathcal{Z}, \tag{6.17}$$

$$\sum_{(\boldsymbol{z},u) \in \mathcal{K}} \alpha(\boldsymbol{z},u)\, r(\boldsymbol{z},u) \leq C_{\text{total}},$$

$$\alpha(\boldsymbol{z},u) \geq 0 \qquad \forall (\boldsymbol{z},u) \in \mathcal{K}.$$

This linear program is called the primal problem.

In [48], it is shown that an optimal stationary strategy π_g^* of the constrained optimization problem (6.12) exists if and only if the primal problem (6.17) has a solution $\boldsymbol{\alpha}^* = \{\alpha^*(\boldsymbol{z},u) | (\boldsymbol{z},u) \in \mathcal{K}\}$. Moreover, an optimal solution of (6.17) uniquely determines an optimal stationary strategy $\pi_g^* \in \Pi_g(S)$. An optimal stationary strategy π_g^*, thus selects action $u \in \mathcal{C}$ at state $\boldsymbol{z} \in \mathcal{Z}$ with probability

$$\pi_g^*(\boldsymbol{z},u) = \frac{\alpha^*(\boldsymbol{z},u)}{\sum_{u \in \mathcal{C}} \alpha^*(\boldsymbol{z},u)}. \tag{6.18}$$

We should point out that an optimal strategy devised by (6.18) is not necessarily a deterministic strategy, in contrast to a strategy that minimizes the cost function (6.1)

without limitations.

Depending on the utilized numerical method, the computational complexity of finding a solution for the linear program in (6.17) varies. It is known that the complexity of the interior-point method increases polynomially with the number of states in $\mathcal{K}$, where the exponent of the complexity polynomial is not large [49]. Moreover, it is known that the number of iterations required for the numerical method to converge is in the order of $O(\log(1/\epsilon))$, where ϵ is the accuracy of the outcome of the numerical method. Here, the size of $\mathcal{K}$ increases exponentially with the number of genes n in the PBN model with control. The goal, in the application of interest, is not to model fine-grained molecular interactions among a host of genes, but rather to model a limited number of genes, typically with very coarse quantization, whose regulatory activities are significantly related to a particular aspect of a specific disease. Hence, the proposed method is easily up to the task of handling the limited size networks with which we are dealing.

B. Constrained Intervention in a Mutated Mammalian Cell Cycle Probabilistic Boolean Network

In this section, we utilize the context-sensitive PBN for the mutated mammalian cell cycle regulation proposed in Section III.B. This context-sensitive PBN postulates the mammalian cell cycle with a mutated phenotype. Our proposed constrained intervention method is then applied with various bounds on the frequency of applying treatments; the therapeutic intervention seeks to hinder cell growth in the absence of growth factors.

Preventing the states with simultaneously down-regulated *CycD* and *Rb* is the objective of intervention. In a devised constrained intervention strategy, if the control

is high, $u = 1$, then the state of the control gene is reversed; if $u = 0$, then the state of the control gene remains unchanged. The control gene can be any of the genes in the model except *CycD*.

We assume that the cost of the states with down-regulated *Rb* and *CycD* is higher than those for the states in which these two genes are not simultaneously down-regulated. We also consider the cost of applying a control action, which increases the cost of each state. We postulate the following cost-per-stage function:

$$c(\boldsymbol{z}, u) = \begin{cases} 0, & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{D} \\ 9, & \text{if } u = 0 \text{ and } \boldsymbol{z} \in \mathcal{U} \\ 1, & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{D} \\ 10, & \text{if } u = 1 \text{ and } \boldsymbol{z} \in \mathcal{U}, \end{cases} \tag{6.19}$$

where $\mathcal{U}$ and $\mathcal{D}$ are the sets of undesirable and desirable states, respectively. A state $\boldsymbol{z}$ is desirable, i.e. belongs to $\mathcal{D}$, if (*CycD*,*Rb*) $\neq (0, 0)$; and undesirable, i.e. belongs to $\mathcal{U}$, if (*CycD*,*Rb*) $= (0, 0)$. We select an arbitrary objective cost-per-stage; however, the cost function is selected so that applying the control to prevent the undesirable states is preferable in comparison to not applying control and remaining in an undesirable state. Assuming the preceding cost-per-stage function, we can compute intervention strategies for the context-sensitive PBN associated to the cell-cycle network according to various constraints.

Figure 17 depicts the steady-state distribution of the states when there is no intervention. According to this figure, the aggregated probability of the states with simultaneously down-regulated *CycD* and *Rb* is close to 0.2. In other words, the model predicts that the mutated cell-cycle will be in the cancerous states nearly 20% of its time in the long run.

Similar to (4.11), we define ΔP_g to be the percentage change in the aggregated

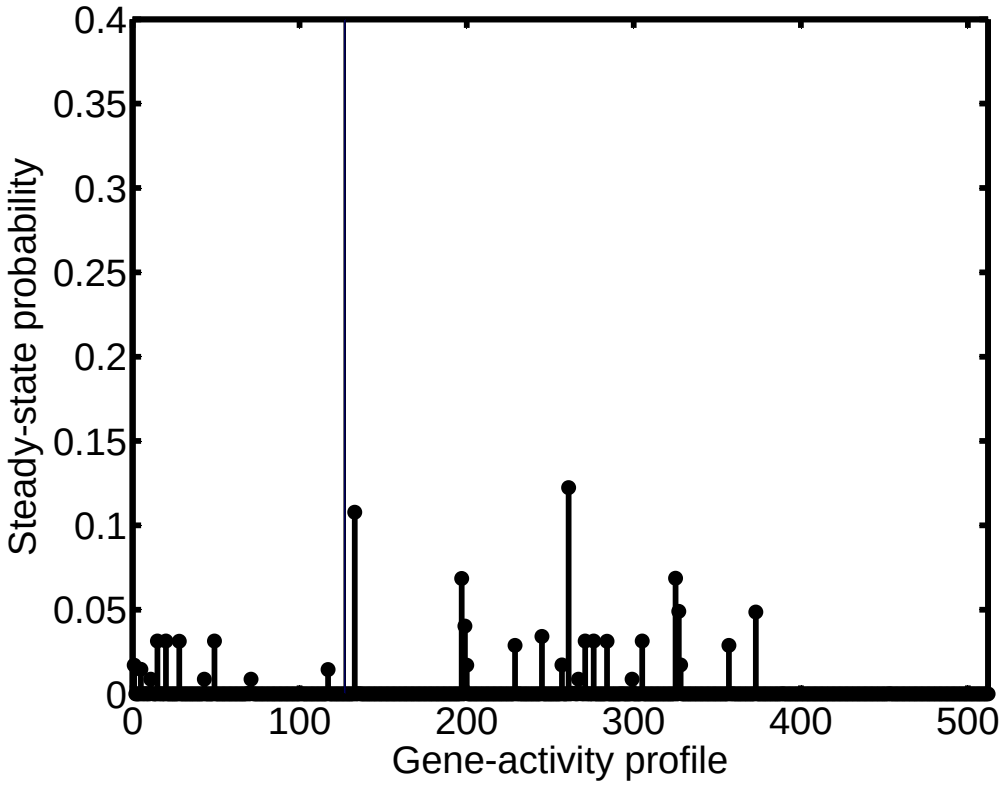


Fig. 17. The steady-state probability of states of the context-sensitive PBN associated with the mammalian cell-cycle network before intervention. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable ones in $\mathcal{D}$.

probability of undesirable states with simultaneously down-regulated *CycD* and *Rb* with and without intervention when gene g is selected as the control gene. As a performance measure, ΔP_g indicates the percentage of the reduction in the likelihood of cancerous situations in the long run.

If we assume that we can alter the expression level of any gene in the network as a therapeutic method, then it is natural to ask which gene should be used to alter the behavior of the model. To this end, we find a constrained intervention strategy for each gene in the network using the intervention method explained in Section VI.A, while limiting the expected number of times treatment can be applied. First, we

assume that the PBN's initial state is the undesirable state in $\mathcal{U}$ with the highest probability in the steady-state distribution of states prior to intervention. Table VIII lists the value of ΔP_g corresponding to each control gene in the network. Here, we vary the upper bound on the frequency of applying intervention and find the corresponding constrained strategies.

Table VIII. The ΔP_g for the intervention strategy based on various control genes and various constraint bounds.

Control Gene g	C_{total}									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Rb	61.96	98.32	98.33	98.33	98.33	98.33	98.34	98.34	98.34	98.34
E2F	57.43	97.36	98.00	98.00	98.00	98.01	98.01	98.02	98.02	98.02
CycE	28.37	28.41	28.41	28.44	28.44	28.46	28.46	28.47	28.49	28.51
CycA	16.56	16.59	16.60	16.61	16.62	16.64	16.65	16.65	16.69	16.69
*Cdc*20	39.15	41.44	41.47	41.48	41.48	41.50	41.51	41.52	41.53	41.61
*Cdh*1	27.55	40.58	41.51	41.56	41.56	41.57	41.62	41.62	61.63	41.65
*UbcH*10	6.49	6.50	6.52	6.56	6.57	6.59	6.61	6.64	6.66	6.69
CycB	39.33	41.85	41.86	41.89	41.91	41.92	41.92	41.96	41.99	41.99

Among all the genes, *Rb* offers the best performance when control can be applied without any constraint, based strictly on minimization of the objective cost function, $C_{\text{total}} = 1$. After applying the unconstrained control strategy designed for *Rb*, the aggregated probability of undesirable states is significantly altered (Figure 18). To avoid the undesirable states in $\mathcal{U}$, we utilize the intervention strategy devised by the proposed method in Section VI.A for the case when there is no bound on the expected number of treatments. In this scenario, let us assume that the state at a decision epoch indicates that $CycD = 0$, $Rb = 1$, $E2F = 1$, $CycE = 1$, $CycA = 0$, $Cdc20 = 0$, $Cdh1 = 1$, $UbcH10 = 0$, and $CycB = 0$. The devised stationary intervention strategy, which is a mapping from the states to the action set $\mathcal{C}$, indicates

that, for the observed state, the value of control gene Rb should be toggled with probability one. Consequently, we should use an appropriate inhibitor to forcefully down regulate the control gene Rb. Hence, the state would be forced from state $(0, 0, 1, 1, 0, 0, 1, 0, 0)$ to state $(0, 1, 1, 1, 0, 0, 1, 0, 0)$ after this intervention. Although the techniques to implement such a strategy, i.e. effectively altering the expression of gene Rb, using its enhancers and inhibitors may not be fully understood within the domain of current medical practice, almost surely these techniques will have detrimental side-effects. The constrained stationary intervention designed by the proposed procedure enables us to restrict the expected number of interventions a patient may receive during therapy. Hence, we could accordingly adjust our intervention strategy when the side-effects of drugs effecting the regulation of gene Rb are known.

Figure 19 indicates that, by using a constrained stationary intervention strategy for the control gene Rb, we can reduce the aggregated probability of the undesirable states to less than 12%, while restricting the number of interventions to at most 10%. We could translate this to restricting the dose of prescribed drugs once knowledge of their side-effects is available. If we only wish to limit the expected number of applied interventions to less than 20%, then we can reduce the chance of the cancerous states by 98%.

According to Table VIII, intervention strategies based on gene $E2F$ performs almost as well as Rb when the constraint is not too tight, $C_{\text{total}} \geq 0.2$. This suggests that, given the side effects of treatments, we may need to consider alternative control genes. The steady-state probability distributions of states after intervention based on $E2F$ are presented in Figure 20 and Figure 21.

Comparing Figure 18 and Figure 20, one can observe that although the final performances of intervening based on these two genes are close, the probability mass of the most probable states after intervention with Rb differs from the one in $E2F$-based

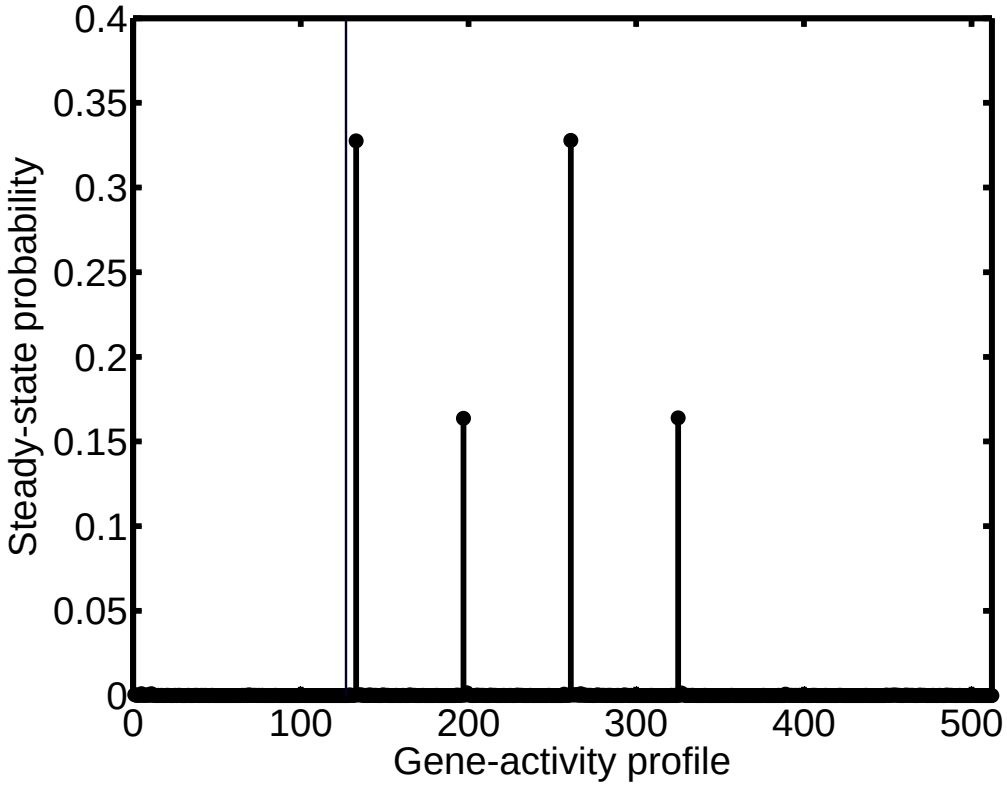


Fig. 18. The steady-state probability of states of the context-sensitive PBN associated with the mammalian cell-cycle network after intervention using Rb as the control gene, when the frequency of applying control is unconstrained, $C_{\text{total}} = 1.0$. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable ones in $\mathcal{D}$.

intervention. This observation suggests that one should utilize systematic analysis along with experimental studies to obtain more effective lever points.

The results of Table VIII indicate that some genes are more sensitive to the bound on the frequency of control. Relaxing the constraint will not improve the result of intervention when the gene *UbcH*10 is selected as the control gene. It is simply not an effective lever point. Genes *CycB* and *Cdc*20 perform relatively well for tightly constrained intervention strategies, but relaxing the limitation on the expected number of treatments does not significantly improve the performance of the strategies

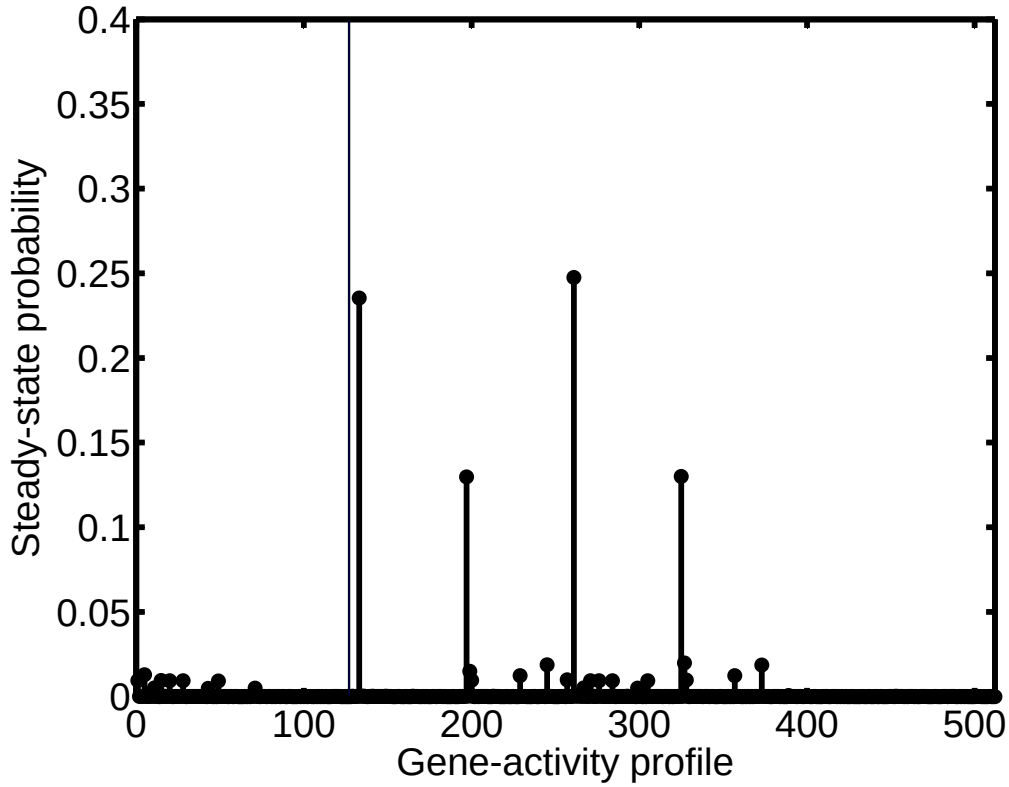


Fig. 19. The steady-state probability of states of the context-sensitive PBN associated with the mammalian cell-cycle network after intervention using Rb as the control gene, when the frequency of applying control is upper bounded by $C_{\text{total}} = 0.1$. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable ones in $\mathcal{D}$.

based on these genes.

Furthermore, if we do not assume that the PBN's initial state is the undesirable state with the highest probability in the steady-state distribution of states prior to intervention, but instead initialize the PBN from an arbitrary undesirable state, we observe that the strategies are robust to the initial state unless the constraint is too tight. For $C_{\text{total}} \geq 0.2$, the values of ΔP do not alter significantly; the performance of the intervention strategy varies more for different initial states when the constraint is tight, $C_{\text{total}} = 0.1$.

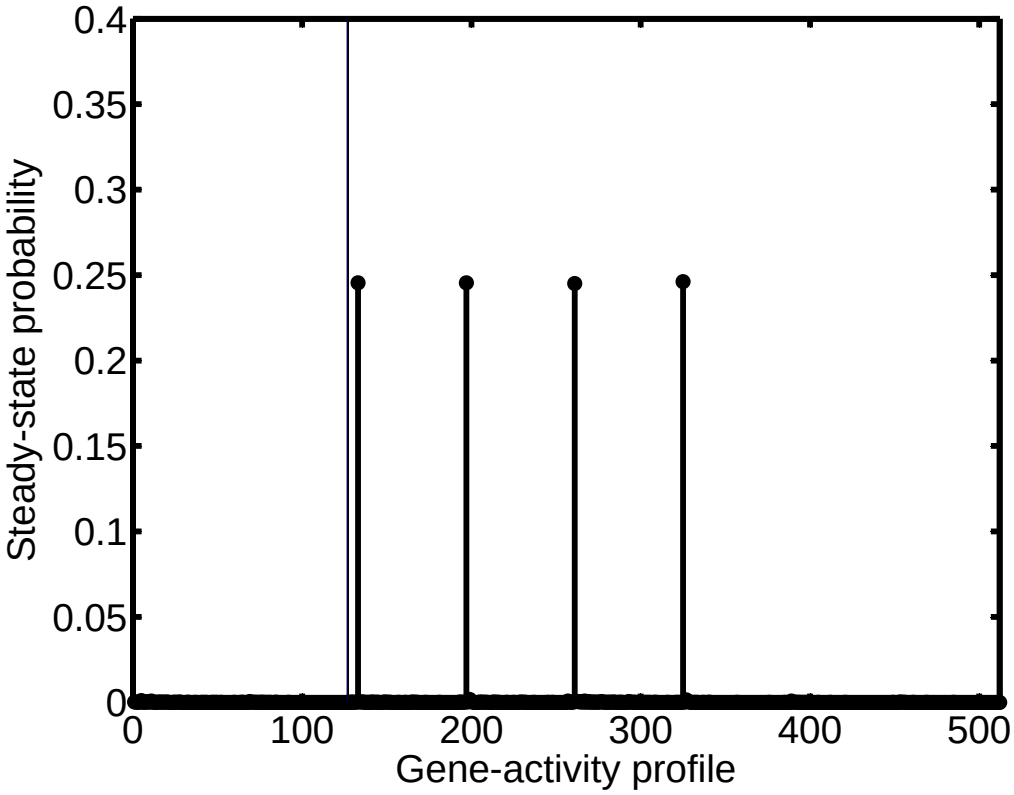


Fig. 20. The steady-state probability of states of the context-sensitive PBN associated with the mammalian cell-cycle network after intervention using E2F as the control gene, when the frequency of applying control is unconstrained, $C_{\text{total}} = 1.0$. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable ones in $\mathcal{D}$.

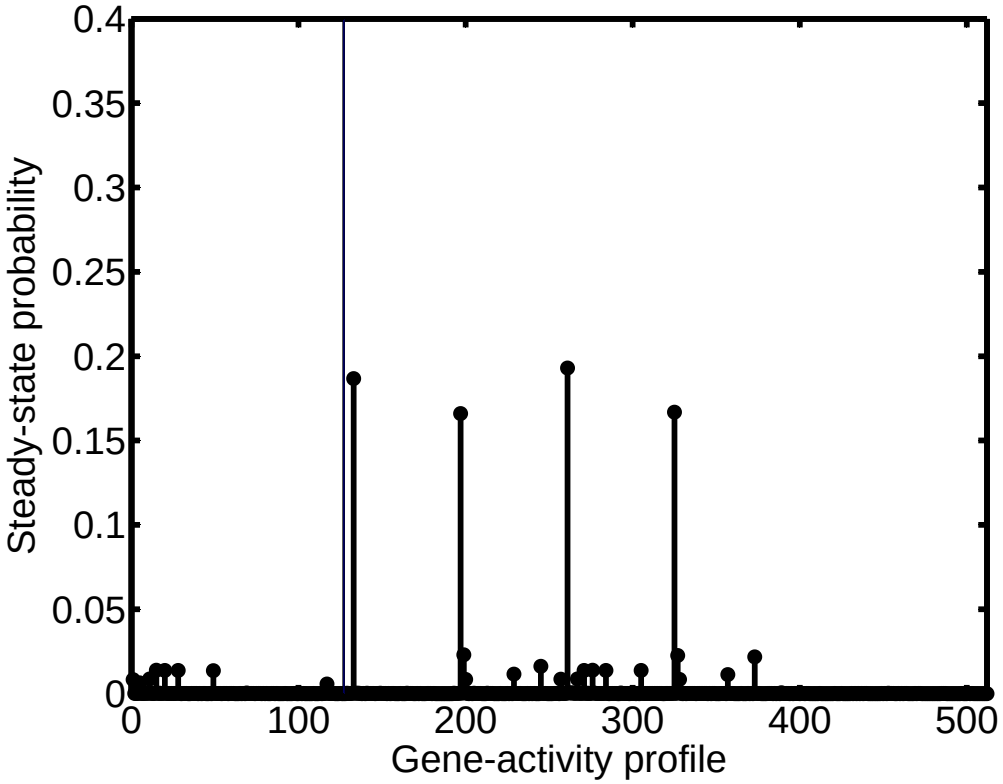


Fig. 21. The steady-state probability of states of the context-sensitive PBN associated with the mammalian cell-cycle network after intervention using E2F as the control gene, when the frequency of applying control is upper bounded by $C_{\text{total}} = 0.1$. The vertical line separates the undesirable states in $\mathcal{U}$ from the desirable ones in $\mathcal{D}$.

CHAPTER VII

MODEL-FREE INTERVENTION IN MARKOVIAN REGULATORY NETWORKS*

The classical intervention described in Chapter IV requires exact optimization of the cost function (4.5). The effectiveness of a strategy devised from the solution of the optimization (4.6) depends on the accuracy of the underlying regulatory network. Moreover, the computational complexity of optimization problem (4.6) increases exponentially with the number of genes in the model. To mitigate this numerical challenge and bypass the impediment of model estimation, a heuristic method is proposed in this chapter. In the framework of Markovian regulatory networks, this heuristic intervention learns effective strategies from a statistics of a pathological cellular behavior using a *reinforcement learning* scheme [42].

Earlier works reveal the limitation of Markovian decision processes in large networks. In [50], it is noted that designing an intervention strategy for the original network with ten genes is beyond their available computational capacity. Thereafter, they resorted to a reduced model with seven genes. Owing to the same limitation, the numerical experiments in [50] and [6] are also designed for a seven-gene instantaneously random PBN similar to the one presented in Chapter III.

Subsequently, it has been demonstrated that the finite-horizon decision making in a PBN is an NP-hard problem [51]. In general, it is well-known that the direct application of Markov decision making methods is limited by the size of the state-space; that is known as the curse of dimensionality [46]. More precisely, It is known that

the complexity of these kinds of sequential decision making methods increases exponentially with the cardinality of model state-space. Hence, the intervention methods such as the ones proposed in [6] and [7] are only applicable to networks with small state-spaces. To address this complexity issue, we propose a heuristic method based on a reinforcement learning algorithm.

Given the cost structure, the *reinforcement intervention* learns an effective strategy based on several generated trajectories of states and applied actions. These empirical measurements are used to estimate the average total cost with respect to various actions and observed states. The reinforcement intervention yields an effective therapeutic strategy, while possessing constant complexity with respect to the number of genes [42]. It has been shown that in the scenario where a large number of measurements is available the reinforcement intervention strategy converges to the strategy devised by the classical intervention.

A salient feature of the reinforcement intervention is that it is *model-free*, i.e. it does not require perfect knowledge of the model parameters. We should point out that it is still assumed that the dynamics of the underlying system are modeled as a Markovian process. The term model-free implies that it is not required to estimate the parameters of the underlying PBN explicitly. As explained in the previous chapters, the intervention methods proposed so far are model dependent, requiring at least knowledge of the transition probability matrix associated with the underlying Markovian regulatory network. This can be derived from the PBN if the latter is known. Since in practice PBNs are not known except via system identification from experimental data, one is faced with a difficult inference problem [17]. This problem can be avoided by directly inferring the transition probability matrix; however, this is still a formidable task because the complexity of estimating the transition probability distributions of a Markov chain increases exponentially with the number of

genes in the model. When time-course measurements are available, a reinforcement intervention strategy can be implemented directly from the empirical measurements. This intervention method has low complexity, is robust to modeling errors, and is also adaptive to changes in the underlying biological system.

The reinforcement intervention is employed to control the $Wnt5a$-related network described in Section III.A. The performances of the heuristic intervention is compared to that of the classical intervention. As noted in Chapter III, down-regulation of $Wnt5a$ is a reasonable objective for an intervention strategy. A reinforcement intervention strategy is applied to the inferred instantaneously random PBN in Section III.A.

A. Reinforcement Intervention in Markovian Regulatory Networks

If the system and cost structure can be simulated, then it is possible to use repeated simulations to calculate approximate transition probability distributions and an expected immediate cost. Thereafter, classical intervention can be applied to find an optimal intervention strategy.

In the absence of time-course measurements, and for our numerical study we assume that the distributions governing the PBN, the switching probability, the perturbation probability and the probability distribution of selecting constituent networks, are known. However, this assumption is not necessary in practice whenever time-series data set are available.

The complexity of estimating the transition probability distributions and the complexity of classical intervention exponentially increase as the number of genes increases. If we contemplate approximation to reduce the complexity, then reinforcement intervention can be used. Given the time-series measurements, the reinforce-

ment intervention progressively computes an approximation of the value function in Equation (4.6) by observing several sample trajectories of the regulatory network and their associated costs. Hence, it eliminates the computational complexity associated with the explicit estimation of the transition probability distributions.

The fixed-point of Bellman optimality equation is achieved at an optimal strategy of classical intervention, so for each $\boldsymbol{z} \in \mathcal{Z}$ the optimal value function is the solution of Equation (4.7) that is

$$J^*(\boldsymbol{z}) = \min_{u \in \mathcal{C}} \left[\overline{c}(\boldsymbol{z}, u) + \lambda \sum_{\boldsymbol{z}' \in \mathcal{Z}} P_{\boldsymbol{z}}(\boldsymbol{z}'; u) J^*(\boldsymbol{z}') \right]. \tag{7.1}$$

The classical intervention iteratively apply the transformation derived by the Bellman optimality equation to each element of the value function until a fixed-point of (7.1) is found [46]. Therefore, the complexity of value iteration is exponential in the number of genes in the PBN. The computational complexity of each iteration of the value iteration is $\mathbf{O}(2^{2n})$, with respect to the number of genes in the network n.

Using the definition of expected immediate cost in expression (4.4), the Bellman optimality equation (7.1) can be rewritten as

$$J^*(\boldsymbol{z}) = \min_{u \in \mathcal{C}} \left[\sum_{\boldsymbol{z}' \in \mathcal{Z}} P_{\boldsymbol{z}}(\boldsymbol{z}'; u) \left(c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \, J^*(\boldsymbol{z}') \right) \right], \tag{7.2}$$

for each state $\boldsymbol{z} \in \mathcal{Z}$. Accordingly, we can define the *Q-factor* for each state-action pair $(\boldsymbol{z}, u) \in \mathcal{Z} \times \mathcal{C}$ by

$$Q(\boldsymbol{z}, u) \stackrel{\triangle}{=} \sum_{\boldsymbol{z}' \in \mathcal{Z}} P_{\boldsymbol{z}}(\boldsymbol{z}'; u) \left(c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \, J^*(\boldsymbol{z}') \right). \tag{7.3}$$

The relation between the optimal value function of a state $\boldsymbol{z}$ and the Q-factors of the same state is given by

$$J^*(\boldsymbol{z}) = \min_{u \in \mathcal{C}} Q(\boldsymbol{z}, u). \tag{7.4}$$

To compute the Q-factor iteratively, the Bellman optimality equation can be written for the Q-factor as

$$Q(\boldsymbol{z}, u) = \sum_{\boldsymbol{z}' \in \mathcal{Z}} P_{\boldsymbol{z}}(\boldsymbol{z}'; u) \left[c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \min_{u' \in \mathcal{C}} Q(\boldsymbol{z}', u') \right]. \tag{7.5}$$

Using Equation (7.5), one can find the solution of the classical intervention with Algorithm 1, in which the value function is replaced by the Q-factor vector.

Algorithm 1 Q-factor version of the classical intervention

$k \leftarrow 0$

Setting $\epsilon > 0$

Selecting an arbitrary initial Q-factor vector, Q^0

repeat

$k \leftarrow k + 1$

Q-factors updating: Computing the transformation (7.5) for all $\boldsymbol{z} \in \mathcal{Z}$.

$$Q^{(k)}(\boldsymbol{z}, u) \leftarrow \sum_{\boldsymbol{z}' \in \mathcal{Z}} P_{\boldsymbol{z}}(\boldsymbol{z}'; u) \left[c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \min_{u' \in \mathcal{C}} Q^{(k-1)}(\boldsymbol{z}', u') \right]$$

Checking the convergence: Computing the following for all $\boldsymbol{z} \in \mathcal{Z}$

$$\begin{aligned} J^{(k)}(\boldsymbol{z}) &= \min_{u \in \mathcal{C}} Q^{(k)}(\boldsymbol{z}, u) \\ J^{(k-1)}(\boldsymbol{z}) &= \min_{u \in \mathcal{C}} Q^{(k-1)}(\boldsymbol{z}, u) \end{aligned}$$

until $\| J^{(k)} - J^{(k-1)} \|_\infty < \epsilon$

Finding an optimal strategy: Choose the intervention strategy for all $\boldsymbol{z} \in \mathcal{Z}$

$$\mu_g(\boldsymbol{z}) = \arg \min_{u \in \mathcal{C}} Q^{(k)}(\boldsymbol{z}, u)$$

Estimation of the Q-factor is the objective of Algorithm 1, in which a Q-factor

is an average of a random variable $\Psi(\boldsymbol{z}, u)$. The equation (7.3) can be expressed as

$$\begin{aligned} Q(\boldsymbol{z}, u) &= E\left[c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \min_{u' \in \mathcal{C}} Q(\boldsymbol{z}', u')\right] \\ &= E\left[\Psi(\boldsymbol{z}, u)\right]. \end{aligned} \tag{7.6}$$

If the samples of $\Psi(\boldsymbol{z}, u)$ are generated within the system's simulator or obtained from time-series measurements, then one can estimate its expected value. Let

$$\overline{\Psi}_k(\boldsymbol{z}, u) = \frac{\sum_{i=1}^{k} \psi_i}{k} \tag{7.7}$$

denote the time average estimation of the ensemble average $E\left[\Psi(\boldsymbol{z}, u)\right]$ using k samples, where ψ_i is the ith sample of the random variable $\Psi(\boldsymbol{z}, u)$. Upon a new observation of $\Psi(\boldsymbol{z}, u)$, the value of $\overline{\Psi}_k(\boldsymbol{z}, u)$ can be updated by

$$\overline{\Psi}_{k+1}(\boldsymbol{z}, u) = \overline{\Psi}_k(\boldsymbol{z}, u) - \frac{\overline{\Psi}_k(\boldsymbol{z}, u)}{k+1} + \frac{\psi_{k+1}}{k+1}. \tag{7.8}$$

If $\alpha^{k+1} = \frac{1}{k+1}$, then we have

$$\overline{\Psi}_{k+1}(\boldsymbol{z}, u) = \overline{\Psi}_k(\boldsymbol{z}, u)\left(1 - \alpha^{k+1}\right) + \alpha^{k+1}\psi_{k+1}. \tag{7.9}$$

Hence, given a new system observation, the Q-factor is iteratively updated for the specific state-action pair $(\boldsymbol{z}, u) \in \mathcal{Z} \times \mathcal{C}$ according to the transformation

$$Q^{(k+1)}(\boldsymbol{z}, u) \leftarrow (1-\alpha)Q^{(k)}(\boldsymbol{z}, u) + \alpha\left[c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \min_{u' \in \mathcal{C}} Q^{(k)}(\boldsymbol{z}', u')\right]. \tag{7.10}$$

In general, the revised value iteration algorithm in which the Q-factors are updated according to (7.10) is called the *Q-learning algorithm* [52]. Since the transformation (7.10) is independent of the transition probability distributions of the system, the reinforcement intervention is a model-free method. In reinforcement intervention, the value of the Q-factor for a state-action pair $(\boldsymbol{z}, u)$ is updated whenever a transition

from state $\boldsymbol{z}$ to state $\boldsymbol{z}'$ occurs in the system's simulator or observed in the time-course data set, given the action u is selected randomly among all the possible actions $\mathcal{C}$.

In the reinforcement intervention, the time-course measurements generate trajectories of state-action pairs; hence, some Q-factors may be updated more often than others. An appropriate step-size α_k is needed to guarantee the convergence of the reinforcement intervention to an optimal intervention strategy despite the asynchronous updating of the Q-factors [52]. Several step-sizes are proposed with the general form $\frac{C}{a+k}$, where C and a can be any positive constants [52]. As a general rule, the step-size should be small, and diminish to zero at a suitable rate [53].

Here, we assume a simple form for the step-size inspired mainly by the argument that the ensemble average can be estimated using the time average of the sample data. We denote the number of times the state-action pair $(\boldsymbol{z}, u)$ is visited by $v(\boldsymbol{z}, u)$. We use $v(\boldsymbol{z}, u)$ to define the step-size α_k equal to $\frac{C}{k}$, where C is a positive constant in the interval $(0, 1)$ and k is $v(\boldsymbol{z}, u)$, whenever the state-action pair $(\boldsymbol{z}, u)$ occurs.

The reinforcement intervention is summarized as Algorithm 2. The complexity of each iteration in Algorithm 2 is $O(1)$ with respect to the number of genes in the network n. Hence, the reinforcement intervention runs in polynomial time complexity with respect to the number of genes in the network.

Moreover, the reinforcement intervention reduces the memory complexity of the classical intervention. In Algorithm 2, the values of the Q-factors are stored explicitly in a tabular form. The algorithm requires $O(2^n)$ memory units; whereas in the classical intervention the required memory is $O(2^{2n})$ memory units. This latter quantity stems from the fact that we must store 2^n values of the value function for 2 actions along with 2^{2n} entries of the transition probability matrices at each iteration of the algorithm. Hence, the required memory of the classical intervention $O(2^{2n} + 2 \times 2^n)$ has a growth of $O(2^{2n})$. Consequently, the memory complexity is considerably reduced

using the reinforcement intervention instead of the classical intervention. The high memory complexity of the classical intervention contributes to its limited applicability to large regulatory networks.

If all the state-action pairs $(\boldsymbol{z}, u)$ are visited infinitely often then for each state-action pair the estimated expected value $\overline{\Psi}_k(\boldsymbol{z}, u)$ converges to its ensemble average $E[\Psi(\boldsymbol{z}, u)]$ with probability one. Hence, we expect that a heuristic strategy computed by the reinforcement intervention converges to the optimal strategy devised by the classical intervention. The convergence of a heuristic strategy to the optimal one is proved in [53] for the general Q-learning method, and our numerical results in Section VII.B support this fact for our case study. In other words, the learning duration of the reinforcement intervention should increase as the number of genes in the network increases in order to obtain a heuristic strategy close to an optimal one. Therefore, the reinforcement intervention, as any other learning algorithm, may not be suitable for extremely large networks.

The maximum size of the intervention problem which can be solved by our heuristic method is hardware dependent. For instance, our current hardware configuration (single Xeon processor and one-GB memory) can obtain near optimal intervention strategy within 10^7 learning periods for a synthetic 15-gene regulatory network. Given more memory and processing power, an accurate intervention strategy can be determined for significantly larger networks within reasonable time. Hence, the computation time is not an issue for us. In the application of interest the goal is not to model fine-grained molecular interactions among a host of genes, but rather to model a limited number of genes, typically with very coarse quantization, whose regulatory activities are significantly related to a particular aspect of a specific disease, such as metastasis in melanoma [7]. The proposed reinforcement intervention is easily up to the task of handling the limited size networks with which we are dealing.

B. Reinforcement Intervention in a Metastatic Melanoma Instantaneously Random Probabilistic Boolean Network

In this section, we apply reinforcement intervention to control the instantaneously random PBN related to metastasis in melanoma constructed in Chapter III, and compare the performance of the heuristic strategy to that of exact one devised by the classical intervention. We consider a ten-gene instantaneously random PBN with 1024 states because our objective is to investigate how an approximate strategy performs in comparison to an optimal strategy. Computing optimal strategies for networks beyond ten genes is not practical with our current computational capability.

After quantifying the multivariate relationships of 587 genes among a sample of melanoma patients, [25] constructed a ten-gene regulatory network involving *Wnt*5*a*. Using instantaneously random PBNs consisting of seven of these genes, subsequent studies developed finite-horizon [32] and infinite-horizon [33] intervention strategies. The reduction to seven genes was dictated by computational requirements. In particular, the transition probability distributions of the Markov chain associated with an instantaneously random PBN are required in the classical intervention presented in these studies. Owing to the exponentially increasing complexity of the classical intervention with the number of genes, approximation of a ten-gene instantaneously random PBN with a seven-gene instantaneously random PBN is a way to make the computations feasible. Moreover, the complexity of estimating transition probability distributions restricts the number of constituent networks in the regulatory network. This is the main reason behind using instantaneously random networks in most of the earlier studies in [6].

We consider a ten-gene instantaneously random PBN consisting of *Wnt*5*a*, *pirin*, *S*100*p*, *Ret*1, *Mmp*3, *Phoc*, *Mart*1, *Hadhb*, *Synuclein*, and *Stc*3 as explained in Sec-

tion III.A. The above order of genes is used in the gene-activity profile, with $Wnt5a$ as the most significant bit and $Stc3$ as the least significant bit. This order of genes in the gene-activity profile facilitates the presentation of our results and does not affect the computed intervention strategy.

As before, having the down-regulation of $Wnt5a$ as the objective, we apply the reinforcement intervention of Algorithm 2 to the inferred $Wnt5a$-related instantaneously random PBN. If the action is high, $u = 1$, then the state of gene Pirin is reversed; if $u = 0$, then the state of Pirin remains unchanged. Pirin has been chosen as the control gene to make a fair comparison with the previous studies for down-regulating the expression of $Wnt5a$ [6]; otherwise, as we showed in Chapter IV, $S100p$ is the most effective control gene for the context-sensitive PBN of this case study with identical parameters.

A cost-per-stage $c(\boldsymbol{z}, \boldsymbol{z}', u)$ is used to compare the reinforcement intervention strategy with the classical intervention strategy. It is assumed that the cost-per-stage is higher, if in the successor state $\boldsymbol{z}'$ $Wnt5a$ is up-regulated. It is also assumed that whenever the intervention is applied, given $Wnt5a$ in the successor state remains unchanged, the cost-per-stage is higher in comparison to when the intervention is not applied. The cost are assigned in a way such that applying the intervention to prevent the undesirable gene-activity profiles is preferable in comparison to not applying intervention and transiting to an undesirable gene-activity profile. We postulate the following cost-per-stage function:

$$c(\boldsymbol{z}, \boldsymbol{z}', u) = \begin{cases} 0, & \text{if } u = 0 \text{ and } \boldsymbol{z}' \in \mathcal{D} \\ 5, & \text{if } u = 0 \text{ and } \boldsymbol{z}' \in \mathcal{U} \\ 1, & \text{if } u = 1 \text{ and } \boldsymbol{z}' \in \mathcal{D} \\ 6, & \text{if } u = 1 \text{ and } \boldsymbol{z}' \in \mathcal{U}, \end{cases} \tag{7.11}$$

where $\mathcal{U}$ and $\mathcal{D}$ are the sets of undesirable and desirable states, respectively. A state $\boldsymbol{z}$ is desirable, i.e. belong to $\mathcal{D}$, if $Wnt5a = 0$; and undesirable, i.e. belong to $\mathcal{U}$, if $Wnt5a = 1$. The values of cost-per-stage function have been chosen to be in line with the earlier studies [6]. Assuming the cost-per-stage function (7.11), we compute an optimal intervention strategy for the ten-gene instantaneously random PBN. With four constituent Boolean networks, estimation of the transition probability distributions takes more than three days with our current hardware configuration (single Xeon processor and one-GB memory). In this hardware configuration, going beyond the binary-valued instantaneously random PBN with 1024 states to a ternary-valued instantaneously random PBN with 59049 states or a context-sensitive PBN with 4096 states enormously increases the estimation time of the transition probability distributions.

Figure 22 depicts the steady-state distribution of the gene-activity profiles when there is no intervention. According to Figure 22, the aggregated probability of the gene-activity profiles with up-regulated $Wnt5a$ is higher than the gene-activity profiles with down-regulated $Wnt5a$. Also, the most probable undesirable gene-activity profile $(1, 0, 1, 0, 1, 0, 0, 1, 0, 0)$ has the highest probability.

After intervention in the $Wnt5a$-related instantaneously random PBN based on an optimal intervention strategy, the steady-state distribution of the gene-activity profile is modified. According to Figure 23, the probability of desirable gene-activity profiles with down-regulated $Wnt5a$ is increased, and the probability of the most probable undesirable gene-activity profile $(1, 0, 1, 0, 1, 0, 0, 1, 0, 0)$ is reduced to 0.007. After intervention with an optimal intervention strategy, the most probable gene-activity profile is $(0, 1, 1, 0, 0, 0, 1, 1, 1, 1)$, which has down-regulated $Wnt5a$.

Similar to expression (4.11), we define ΔP to be the percentage of change in the aggregated probability of the gene-activity profiles with $Wnt5a = 1$ before and after

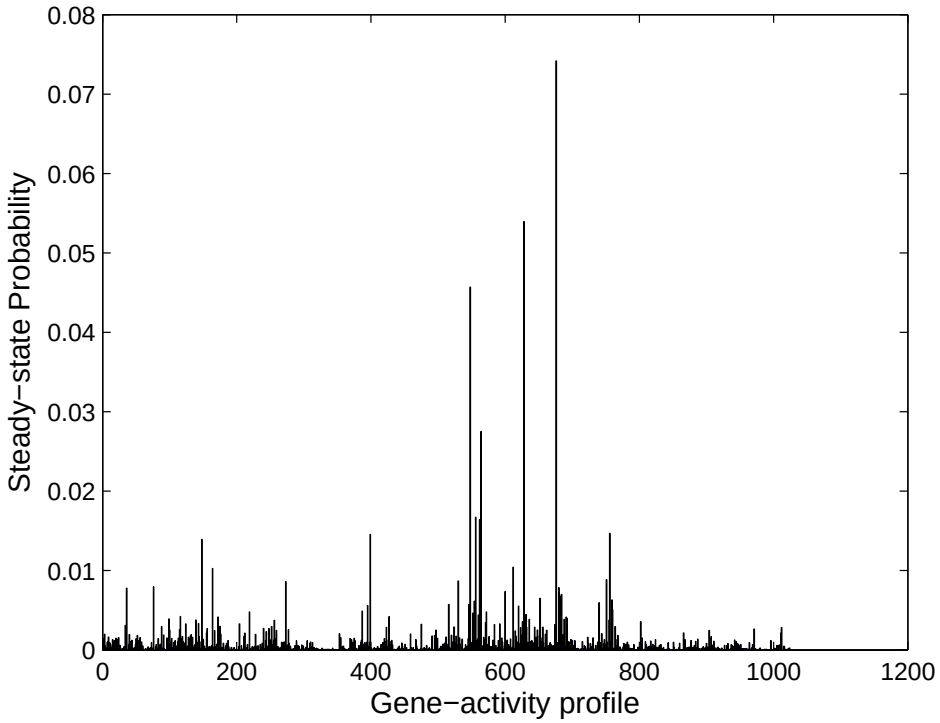


Fig. 22. Steady-state distribution of gene-activity profile of the ten-gene instantaneously random PBN prior to intervention.

the intervention. As a performance measure, ΔP indicates the percentage of reduction in the total probability of the undesirable gene-activity profiles in the steady-state. For an optimal intervention strategy, determined by the classical intervention, we have $\Delta P = 23.1\%$.

In order to compare the reinforcement intervention with the classical intervention, we execute the reinforcement method for different learning durations k_{max}. An approximate intervention strategy is used to find the steady-state distribution of the gene-activity profile.

Figure 24 and Figure 25 show the steady-state distributions of the gene-activity profile when $k_{max} = 10^3$ and $k_{max} = 10^5$, respectively. The computed strategy after only a short learning duration, $k_{max} = 10^3$, does not reduce the likelihood of unde-

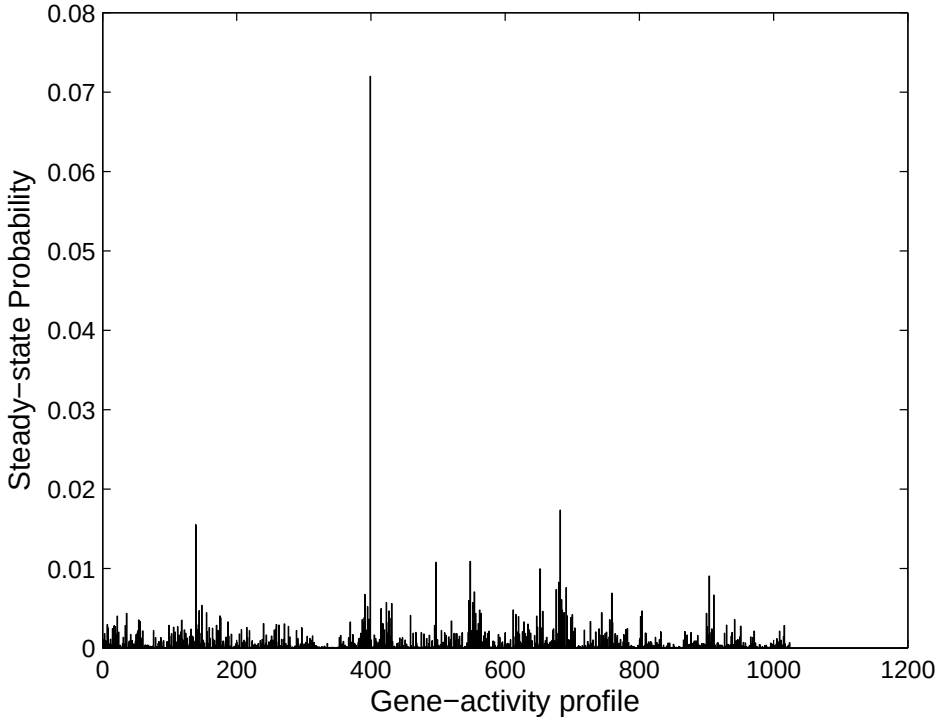


Fig. 23. Steady-state distribution of gene-activity profile after intervention with an optimal intervention strategy.

sirable gene-activity profiles, but when the algorithm's learning duration increases to $k_{max} = 10^5$ the aggregated probability of undesirable gene-activity profiles is reduced.

Comparing Figure 25 and Figure 23, we observe that the probability distributions of the gene-activity profile are similar after intervention with an optimal strategy and a heuristic strategy with long enough learning duration.

Table IX compares the steady-state probabilities of the two most probable gene-activity profiles before and after intervention. The steady-state probability of the gene-activity profile $(1, 0, 1, 0, 1, 0, 0, 1, 0, 0)$ with the highest probability prior to any intervention reduces by almost the same amount when either an optimal intervention strategy or a heuristic intervention strategy computed after a sufficiently long learning duration is used. Moreover, after intervention with either an optimal or a reinforce-

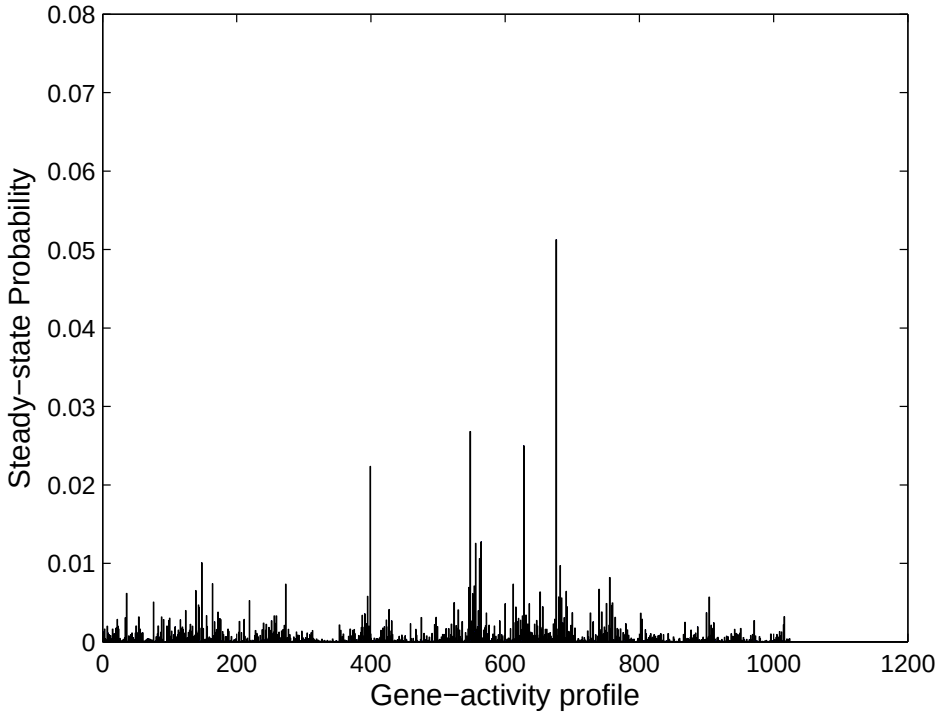


Fig. 24. Steady-state distribution of gene-activity profile after applying an approximate intervention strategy computed by the reinforcement intervention with $k_{max} = 10^3$.

ment intervention strategy, the desirable gene-activity profile $(0, 1, 1, 0, 0, 0, 1, 1, 1, 1)$ has the highest steady-state probability.

As the duration of learning in the reinforcement intervention increases, its performance gets closer to that of the classical intervention. Figure 26 shows the value of ΔP for an optimal intervention strategy, as well as this value for reinforcement intervention strategies derived by the reinforcement method with various learning durations. The performance of the reinforcement intervention converges to the optimal intervention devised by the classical method. We expect to observe this behavior because as the learning duration increases, the estimate of the Q-factor vector becomes more accurate.

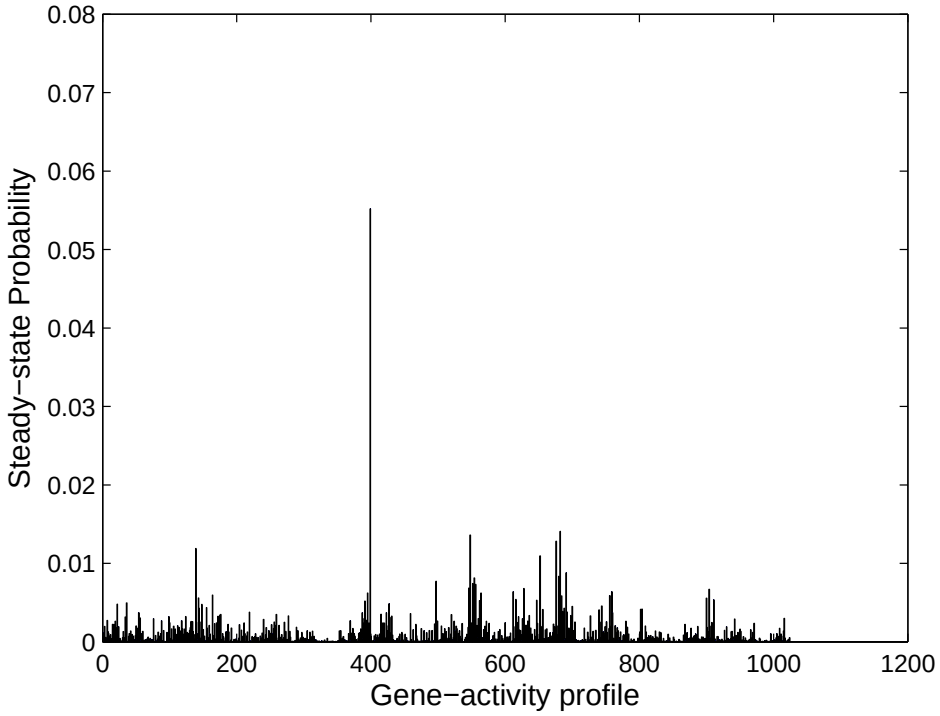


Fig. 25. Steady-state distribution of gene-activity profile after applying an approximate intervention strategy computed by an reinforcement intervention with $k_{max} = 10^5$.

Moreover, we observe that the time-complexity of the reinforcement intervention increases linearly with the number of iterations. Figure 27 depicts the time it takes to run the reinforcement intervention with our current hardware configuration. The execution time of the reinforcement intervention is still tolerable when the learning duration is increased to achieve an acceptable performance for the algorithm. Hence, the melanoma case study reveals that not only the reinforcement intervention does provide near-optimal performance, but considerably reduces the time complexity and the memory complexity of classical intervention. Through selecting an appropriate learning duration, we can have a tradeoff between the desirable accuracy of the approximate intervention strategy and the execution time of the reinforcement

Table IX. Steady-state probability of the most probable gene-activity profile prior and after intervention. The gene-activity profiles $(1,0,1,0,1,0,0,1,0,0)$ and $(0,1,1,0,0,0,1,1,1,1)$ are represented by their binary bijections 676 and 399, respectively.

Gene-expression	No-Cont	Opt-Cont	k_{max}			
			10^3	10^4	10^5	10^7
Max Prob Gene-Exp	676	399	676	399	399	399
Gene-Exp 676 Prob	0.07	0.007	0.05	0.02	0.01	0.009
Gene-Exp 399 Prob	0.01	0.07	0.02	0.04	0.07	0.07

intervention.

C. Reinforcement Intervention Versus Mean First-Passage Time Intervention

The Mean first-passage time intervention (MFPT) is a heuristic method that determines an effective intervention policy for PBN in a model-free fashion [42]. Similar to the reinforcement intervention, the model-free MFPT intervention runs in constant time complexity with respect to the number of genes in the network. Moreover, the learning duration of the MFPT intervention, similar to that of the reinforcement intervention, should increase as the number of genes in the network increases in order to obtain a heuristic strategy close to an optimal strategy.

We compare the performance of the reinforcement and MPFT interventions. To this end, we apply the reinforcement and MFTP strategies to control the instantaneously random PBN related to a melanoma case study inferred in Chapter III, and compare the performance of the two methods to that of an optimal strategy.

We postulate an appropriate cost-per-stage function similar to expression (7.11). The desirable states are assigned lower cost compared to the undesirable states. For

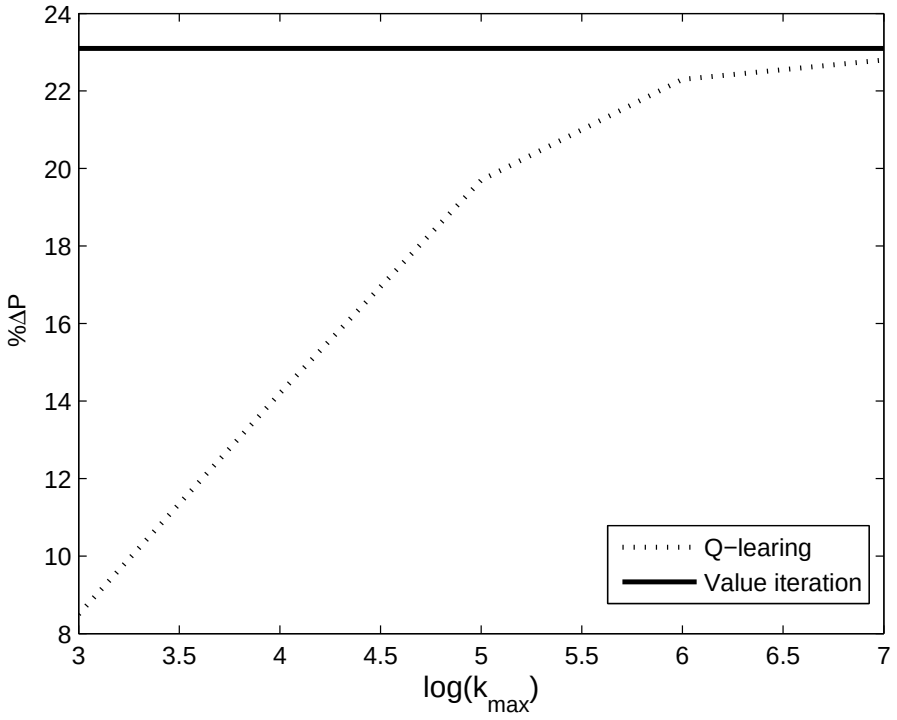


Fig. 26. ΔP of an approximate intervention strategy versus an optimal intervention strategy as a function of logarithm of learning duration.

fair comparison of the two algorithms, the cost of control is considered negligible compared to the cost of undesirable states. Having the down-regulation of $Wnt5a$ as the objective, we apply the reinforcement and MFPT strategies to the inferred instantaneously random PBN. Again, Pirin has been chosen as the control gene.

As a performance measure, ΔP^{opt}, ΔP^{RL}, and ΔP^{MFPT} indicate the percentages of reduction in the total probability of the undesirable states in the steady-state when the classical, reinforcement, and MFPT intervention strategies are applied, respectively. These three performance metrics are computed similar to expression (4.11).

We generate time-course data for 10^6 time-steps from the existing model. We apply reinforcement and MFPT strategies after each 10^k time-steps, for $k = 3, \ldots, 6$.

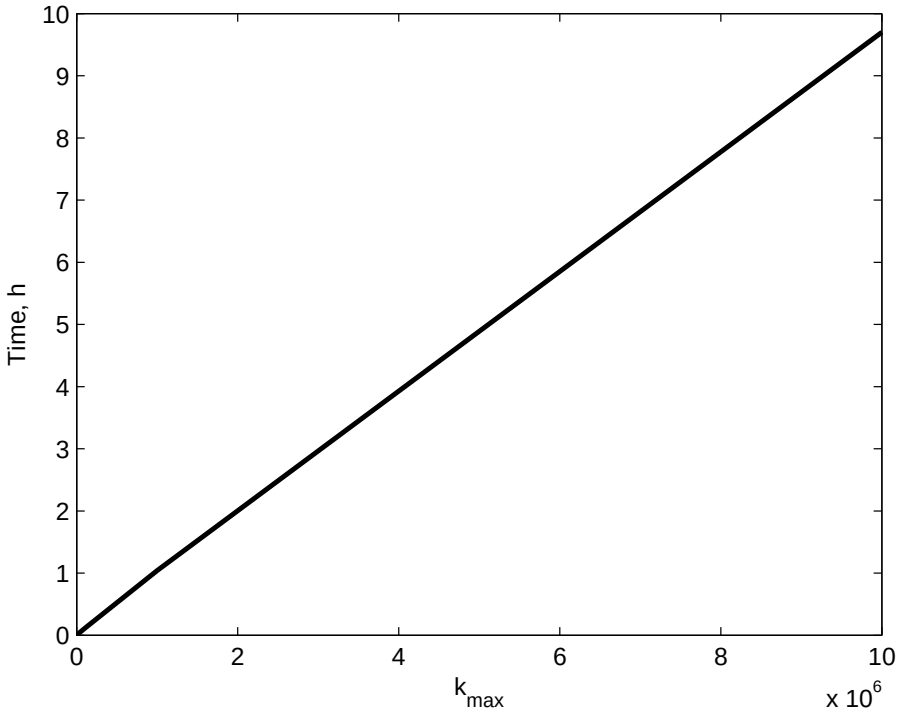


Fig. 27. Execution time of the reinforcement intervention versus learning duration.

Hence, ΔP^{RL} and ΔP^{MFPT} are functions of the learning duration. On the other hand, ΔP^{opt} is computed from the instantaneously random PBN by directly solving the classical intervention optimization.

Figure 28 shows $\Delta P^{\mathrm{opt}} - \Delta P^{\mathrm{RL}}$ and $\Delta P^{\mathrm{opt}} - \Delta P^{\mathrm{MFPT}}$ as a function of the logarithm of the learning duration. After 10^3 time-course data points, $\Delta P^{\mathrm{opt}} - \Delta P^{\mathrm{MFPT}}$ is 0.114 while $\Delta P^{\mathrm{opt}} - \Delta P^{\mathrm{RL}}$ is 0.166. In particular, for lower numbers of observations, which correspond better to feasible experimental conditions, the approximation by the MFPT intervention outperforms the reinforcement intervention. On the other hand, after 10^6 time-course data points, $\Delta P^{\mathrm{opt}} - \Delta P^{\mathrm{MFPT}}$ is 0.003 while $\Delta P^{\mathrm{opt}} - \Delta P^{\mathrm{RL}}$ is 0.002. As the size of the training data increases, the performance of the reinforcement intervention seems to overtake that of the MFPT intervention.

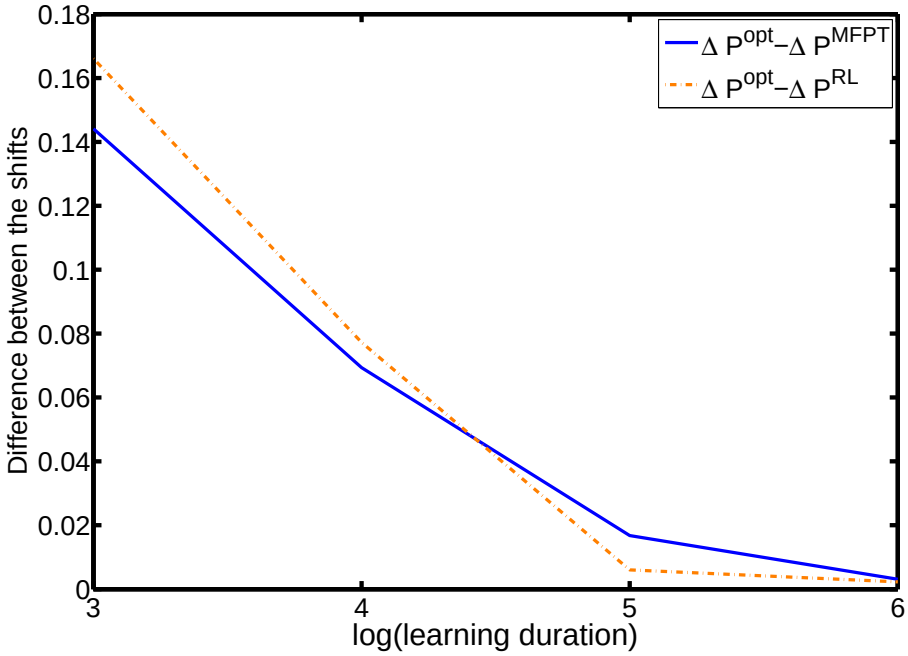


Fig. 28. $\Delta P^{opt} - \Delta P^{RL}$ and $\Delta P^{opt} - \Delta P^{MFPT}$ as a function of the logarithm of the learning duration.

As we mentioned earlier, in the reinforcement intervention, the value of the Q-factor for a state-action pair $(\boldsymbol{z}, u) \in \mathcal{Z} \times \mathcal{C}$ is updated given the action u is selected randomly among all the possible actions. To this end, in the reinforcement intervention, the two possible values of action $u \in \mathcal{C}$ should be applied to the model with equal probability. In contrast, the MFPT intervention does not require any application of the intervention for obtaining the MFPT intervention strategy.

In the presented comparison, the MFPT intervention does not take into account the cost of intervention. Therefore, this comparison is valid when the cost of intervention is negligible compared to the cost of undesirable states. A situation that does not necessarily hold in everyday applications. This comparison suggests that a combination of heuristic intervention methods may achieve a better intervention strategy in a real treatment discovery.

Algorithm 2 Reinforcement Intervention

$Q(\boldsymbol{z}, u) \leftarrow 0$ for all $\boldsymbol{z} \in \mathcal{Z}, u \in \mathcal{C}$

$v(\boldsymbol{z}, u) \leftarrow 0$ for all $\boldsymbol{z} \in \mathcal{Z}, u \in \mathcal{C}$

Setting $0 < C < 1$

Setting k_{max}

Selecting an arbitrary initial state $\boldsymbol{z}$.

for $k = 0$ to k_{max} **do**

Action Selection: Selecting action $u \in \mathcal{C}$ randomly, given the current state is $\boldsymbol{z}$.

Extracting information from time-series data: According to the transition in a time-course measurements of the system's simulator, the successor state $\boldsymbol{z}'$ as well as the earned cost-per-stage $c(\boldsymbol{z}, \boldsymbol{z}', u)$ are determined in a transition from state $\boldsymbol{z}$ to state $\boldsymbol{z}'$ under action u. Hence, the following updates are performed:

$v(\boldsymbol{z}, u) \leftarrow v(\boldsymbol{z}, u) + 1$

$\alpha \leftarrow \frac{C}{v(\boldsymbol{z}, u)}$.

Updating $Q(\boldsymbol{z}, u)$ **:** The value of $Q(\boldsymbol{z}, u)$ is updated according to transformation (7.10).

$$Q(\boldsymbol{z}, u) \leftarrow (1 - \alpha)\, Q(\boldsymbol{z}, u) + \alpha \left[c(\boldsymbol{z}, \boldsymbol{z}', u) + \lambda \min_{u' \in \mathcal{C}} Q(\boldsymbol{z}', u') \right]$$

Setting $\boldsymbol{z} \leftarrow \boldsymbol{z}'$

end for

Finding the heuristic strategy: Choose the heuristic strategy for all $\boldsymbol{z} \in \mathcal{Z}$

$$\mu_g(\boldsymbol{z}) = \arg \min_{u \in \mathcal{C}} Q(\boldsymbol{z}, u)$$

CHAPTER VIII

CONCLUDING REMARKS

Traditionally, drug and treatment discoveries are performed by identifying individual components and studying their specific functionalities, rather than relating molecular components to their systemic functions. Looking at a few aspects of an organism at a time has shown limited success in devising effective treatments for complex diseases such as cancer. The advent of new technologies in recent years has enabled us to catalog the human genome, and has placed us in a position to amend the traditional methods to remedy complex health problems. Now that it is evident that many uncured diseases are almost never caused by a single gene, protein, or biochemical reaction, we should seek to investigate treatments as actions on integrated and interacting networks of biological components.

Regulatory networks that represent such biological systems are highly structured but incredibly complex. We developed systems-based methods that allow us to understand the interplay between an organism's genome and environmental factors as they relate to a disease. This integrative knowledge is crucial to understand how the human body operates, and how we can best cure complex diseases such as cancer.

In this volume, we proposed several algorithmic approaches to design systems-based therapies based on the information delineated in regulatory networks. The proposed schemes aim to overcome engineering issues related to complexity, inference, and robustness, and also aim to develop intervention strategies commensurate with practical medical constraints.

First, we proposed two asynchronous regulatory network frameworks and demonstrated how they can be used to design effective intervention strategies. The DA-PBN model extends the benefits of context-sensitive PBNs by adding the ability to cope

with temporal context as well as structural context. Since asynchronism at the gene level has practical limitations, we introduced the SM-ARN model, in which the asynchronism is at the gene-expression profile level. Empirically measurable timing information of biological systems can be directly incorporated into the SM-ARN model to determine the time-delay distributions between transitions from one gene-expression profile to another. Using the SM-ARN model, we have modeled the dynamics of a mutated mammalian cell cycle regulatory network. The proposed intervention method for the SM-ARN is then used to design a strategy to influence the dynamics of the SM-ARN constructed for the mutated mammalian cell cycle. The goal of the intervention is to reduce the long-run likelihood of undesirable cell growth. The presented numerical studies strongly suggest that our intervention method effectively alters the dynamics of the cell cycle model.

Then, we formulated the constrained intervention method in probabilistic Boolean networks and demonstrated that one can reduce the likelihood of a subset of undesirable states while bounding the expected number of interventions in a therapeutic strategy using the proposed method. We have considered a mutated mammalian cell-cycle network in which the cell growth does not stop in the absence of growth factors. We have then utilized the proposed intervention method to design constrained intervention policies to influence the dynamics of the PBN constructed for the mutated mammalian cell cycle. The goal of intervention is to reduce the chance of undesirable cell proliferation in the long run, while maintaining a bound on the expected number of interventions. The presented numerical studies strongly suggest that constrained intervention can effectively alter the dynamics of the cell-cycle model. Various control genes can be considered given different constraints. The most effective control gene may vary depending on the restrictions imposed on the intervention policies.

At the end, we formulated a model-free algorithm to find an effective interven-

tion strategy. The proposed reinforcement intervention not only lowers computational complexity in comparison to the classical intervention method, it also performs virtually the same as the optimal method when the learning duration is long enough. As shown in the case of a melanoma-related regulatory network, applying the heuristic strategy designed by the reinforcement intervention has the same effect in reducing the likelihood of visiting undesirable gene-expression profiles. The time complexity of the reinforcement intervention is polynomial, whereas the time complexity of the classical intervention is exponential in the number of genes. We can have a trade off between the desirable accuracy of the reinforcement strategy and its computation time. Since the proposed reinforcement intervention is a model-free algorithm, the estimation of transition probabilities for the Markov chain modeling the dynamics of a context-sensitive PBN is not required. Consequently, the proposed method also eliminates the time complexity of estimation processes prior to control, and can further devise an effective intervention strategy directly from time-series measurements.

The proposed systems-based therapeutic methods have only begun to deal with technical and practical issues; much remains to be accomplished relative to these two aims before system-based therapeutic design can be fully integrated into medical practice. There are many directions in which future work can proceed. Following is a few immediate avenues of research.

- **Long-run probability distribution of expression profiles:** Our studies indicate that the current intervention methods, which are derived by minimizing different performance metrics, also reduce the long-run probability of undesirable expression profiles as a byproduct. Alternative intervention design procedures that directly consider the probability of expression profiles as a performance metric should be formulated. Moreover, the outcome of a devised

therapy should produce a biologically viable genotype beside avoiding undesirable cellular phenotypes. This fact is often overlooked in current algorithms, and will be addressed in my future research.

- **Intervention without explicit cost function:** The current methods are intended to reduce the likelihood of the gene-expression profiles associated with aberrant cellular functions. This has been accomplished by avoiding undesirable gene-expression profiles. To this end, cost values are needed to capture the benefits and costs of intervention, and the relative preference of gene-expression profiles. Setting these values in a biologically meaningful way is a daunting task, so it is of practical interest to design intervention methods based on some *partial ordering* of states. I believe that alternative decision making processes can be used to formulate the latter paradigm of intervention design. For instance, given the current medical technology, it would be advantageous to design interventions that eliminate tumors by guiding the trajectory of the diseased cells toward gene-expression profiles that initiate programmed cell-death rather than attempting to restore a particular cellular behavior by avoiding a class of expression profiles. This method only requires a partial ordering of gene-expression profiles, not an explicit definition of the cost function.

- **Protein reporters as the only observable information:** The only observable variables in time-series experiments are a number of reporters. The state of observable variables provide imperfect state information relative to the full state-space of the underlying regulatory system. In this case, the lever points for intervention and their therapeutic effectiveness will depend on a probabilistic mapping between the full state vector and the measurement vector for the observable variables. Since such a mapping is not known apriori, I conjecture

that an adaptive method inspired by the Certainty Equivalence Principle could be utilized in this context; and plan to explore it in the future.

REFERENCES

[1] T. Miyashita and J. C. Reed, "Tumor suppressor *p*53 is a direct transcriptional activator of the human bax gene," *Cell*, vol. 80, no. 2, pp. 193–199, 1995.

[2] L. B. Owen-Schaub, W. Zhang, J. C. Cusack, L. S. Angelo, S. M. Santee, T. Fujiwara, J. A. Roth, A. B. Deisseroth, W. W. Zhang, and E. Kruzel, "Wild-type human *p*53 and a temperature-sensitive mutant induce *Fas/Apo*-1 expression," *Molecular and Cellular Biology*, vol. 15, no. 6, pp. 3032–3040, 1995.

[3] W. S. El-Deiry, T. Tokino, V. E. Velculescu, D. B. Levy, R. Parsons, J. M. Trent, D. Lin, W. E. Mercer, K. W. Kinzler, and B. Vogelstein, "*Waf1*, a potential mediator of *p*53 tumor suppression," *Cell*, vol. 75, no. 4, pp. 817–825, 1993.

[4] S. G. Swisher, J. A. Roth, J. Nemunaitis, D. D. Lawrence, B. L. Kemp, C. H. Carrasco, D. G. Connors, A. K. El-Naggar, F. Fossella, B. S. Glisson, W. K. Hong, F. R. Khuri, J. M. Kurie, J. J. Lee, J. S. Lee, M. Mack, J. A. Merritt, D. M. Nguyen, J. C. Nesbitt, R. Perez-Soler, K. M. W. Pisters, J. B. Putnam, W. R. Richli, M. Savin, D. S. Schrump, D. M. Shin, A. Shulkin, G. L. Walsh, J. Wait, D. Weill, and M. K. A. Waugh, "Adenovirus-mediated *p*53 gene transfer in advanced non-small-cell lung cancer," *Journal of the National Cancer Institute*, vol. 91, no. 9, pp. 763–771, 1999.

[5] M. Bouvet, R. J. Bold, , J. Lee, D. B. Evans, J. L. Abbruzzese, P. J. Chiao, D. J. McConkey, J. Chandra, S. Chada, B. Fang, and J. A. Roth, "Adenovirus-mediated wild-type *p*53 tumor suppressor gene therapy induces apoptosis and suppresses growth of human pancreatic cancer," *Annals of Surgical Oncology*, vol. 5, no. 8, pp. 681–688, 1998.

[6] A. Datta and E. R. Dougherty, *Introduction to Genomic Signal Processing with Control*. Boca Raton, FL, USA: CRC Press, 2006.

[7] I. Shmulevich and E. R. Dougherty, *Genomic Signal Processing*. Princeton, NJ, USA: Princeton University Press, 2007.

[8] M. Bittner, P. Meltzer, Y. Chen, Y. Jiang, E. Seftor, M. Hendrix, 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, V. Sondak, N. Hayward, and J. Trent, "Molecular classification of cutaneous malignant melanoma by gene expression profiling," *Nature*, vol. 406, no. 6795, pp. 536–450, 2000.

[9] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic networks," *Journal of Theoretical Biology*, vol. 22, no. 3, pp. 437–467, 1969.

[10] S. A. Kauffman and S. Levin, "Towards a general theory of adaptive walks on rugged landscapes," *Journal of Theoretical Biology*, vol. 128, no. 1, pp. 11–45, 1987.

[11] S. A. Kauffman, *The Origins of Order: Self-organization and Selection in Evolution*. New York, NY, USA: Oxford University Press, 1993.

[12] N. A. M. Monk, "Oscillatory expression of *Hes*1, *p*53, and *NF-*κ*B* driven by transcriptional time delays," *Current Biology*, vol. 13, no. 16, pp. 1409–1413, 2003.

[13] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, "Probing gene expression in live cells, one protein molecule at a time," *Science*, vol. 311, no. 5767, pp. 1600–1603,

2006.

[14] B. Faryabi, J.-F. Chamberland, G. Vahedi, A. Datta, and E. R. Dougherty, "Optimal intervention in asynchronous genetic regulatory networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 412–423, 2008.

[15] X. Deng, H. Geng, and M. T. Matache, "Dynamics of asynchronous random Boolean networks with asynchrony generated by stochastic processes," *BioSystems*, vol. 88, no. 1-2, pp. 16–34, 2007.

[16] F. Greil and B. Drossel, "The dynamics of critical Kauffman networks under asynchronous stochastic update," *Physical Review Letters*, vol. 95, no. 4 (048701), 2005.

[17] S. Marshall, L. Yu, Y. Xiao, and E. R. Dougherty, "Inference of a probabilistic Boolean network from a single observed temporal sequence," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2007, pp. 32 454–32 569, 2007.

[18] R. Simon and L. Norton, "The Norton-Simon hypothesis: Designing more effective and less toxic chemotherapeutic regimens," *Nature Clinical Practice Oncology*, vol. 3, no. 8, pp. 406–407, 2006.

[19] N. Vasudevan, Y. S. Zhu, S. Daniel, N. Koibuchi, W. W. Chin, and D. Pfaff, "Crosstalk between oestrogen receptors and thyroid hormone receptor isoforms results in differential regulation of the preproenkephalin gene," *Journal of Neuroendocrinol*, vol. 13, no. 9, pp. 779–790, 2001.

[20] Z. Zhang and C. T. Teng, "Estrogen receptor α and estrogen receptor-related receptor α1 compete for binding and coactivator," *Molecular and Cellular Endocrinology*, vol. 172, no. 1-2, pp. 223–233, 2001.

[21] B. Faryabi, J.-F. Chamberland, G. Vahedi, A. Datta, and E. R. Dougherty, "Optimal constrained stationary intervention in gene regulatory networks," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2008, no. 62076, p. 10 pages, 2008.

[22] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang, "Probabilistic Boolean networks: A rule-based uncertainty model for gene regulatory networks," *Bioinformatics*, vol. 18, no. 2, pp. 261–274, 2002.

[23] S. Huang, "Gene expression profiling, genetic networks, and cellular states: An integrating concept for tumorigenesis and drug discovery," *Journal of Molecular Medicine*, vol. 77, no. 6, pp. 469–480, 1999.

[24] R. Thomas, *Kinetic Logic: A Boolean Approach to the Analysis of Complex Regulatory Systems*. Berlin, New York, USA: Springer-Verlag, 1979.

[25] S. Kim, H. Li, E. R. Dougherty, N. W. Cao, Y. D. Chen, M. L. Bittner, and E. B. Suh, "Can Markov chain models mimic biological regulation?" *Journal of Biological Systems*, vol. 10, no. 4, pp. 337–357, 2002.

[26] I. Shmulevich, E. R. Dougherty, and W. Zhang, "From Boolean to probabilistic Boolean networks as models of genetic regulatory networks," *Proceedings of the IEEE*, vol. 90, no. 11, pp. 1778–1792, 2002.

[27] M. Brun, E. R. Dougherty, and I. Shmulevich, "Steady-state probabilities for attractors in probabilistic Boolean networks," *Signal Processing*, vol. 85, no. 10, pp. 1993–2013, 2005.

[28] I. Shmulevich, E. R. Dougherty, and W. Zhang, "Gene Perturbation and Intervention in Probabilistic Boolean Networks," *Bioinformatics*, vol. 18, no. 10, pp.

1319–1331, 2002.

[29] A. Faure, A. Naldi, C. Chaouiya, and D. Theiffry, "Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle," *Bioinformatics*, vol. 22, no. 14, pp. e124–e131, 2006.

[30] M. Chaves, R. Albert, and E. D. Sontag, "Robustness and fragility of Boolean models for genetic regulatory networks," *Journal of Theoretical Biology*, vol. 235, pp. 431–449, 2005.

[31] B. Faryabi, G. Vahedi, J.-F. Chamberland, A. Datta, and E. R. Dougherty, "Intervention in context-sensitive probabilistic boolean networks revisited," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2009, no. 360864, p. 10 pages, 2009.

[32] R. Pal, A. Datta, M. L. Bittner, and E. R. Dougherty, "Intervention in context-sensitive probabilistic Boolean networks," *Bionformatics*, vol. 21, no. 7, pp. 1211–1218, 2005.

[33] R. Pal, A. Datta, and E. R. Dougherty, "Optimal infinite-horizon control for probabilistic Boolean networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2375–2387, 2006.

[34] J. R. Norris, *Markov Chains*. Cambridge, UK: Cambridge University Press, 1998.

[35] I. Harvey and T. Bossomaier, "Time out of joint: Attractors in asynchronous random Boolean networks," in *Proceeding of the 4th European Conference on Artificial Life (ECAL97)*, MIT Press, July 1997, pp. 67–75.

[36] C. Gershenson, "Classification of random Boolean networks," in *Proceedings of the Eighth International Conference on Artificial Life (Artificial Life VIII)*, Sydney, Australia, December 2002, pp. 1–8.

[37] E. A. Di Paolo, "Searching for rhythms in asynchronous Boolean networks," in *Proceedings of the Seventh International Conference on Artificial Life (Artificial Life VII)*, Portland, OR, August 2000, pp. 1–6.

[38] D. Cornforth, D. G. Green, D. Newth, and M. R. Kirley, "Ordered asynchronous processes in natural and artificial systems," in *Proceeding of the 5th Australia-Japan Joint Workshop on Intelligent and Evolutionary Systems*, Dunedin, New Zealand, November 2001, pp. 105–112.

[39] M. Chaves, E. D. Sontag, and R. Albert, "Methods of robustness analysis for Boolean models of gene control networks," *IEE Proceedings in Systems Biology*, vol. 235, pp. 154–167, 2006.

[40] B. Mesot and C. Teuscher, "Critical values in asynchronous random Boolean networks," in *Proceeding of the 7th European Conference on Artificial Life(ECAL03)*, MIT Press, September 2003, pp. 367–377.

[41] R. G. Gallager, *Discrete Stochastic Processes*. Norwell, MA, USA: Kluwer Academic Publisher, 1995.

[42] B. Faryabi, A. Datta, and E. R. Dougherty, "On approximate stochastic control in genetic regulatory networks," *IET Journal of Systems Biology*, vol. 1, no. 6, pp. 361–368, 2007.

[43] A. T. Weeraratna, Y. Jiang, G. Hostetter, K. Rosenblatt, P. Duray, M. Bittner, and J. M. Trent, "*Wnt*5*a* signaling directly affects cell motility and invasion of

metastatic melanoma," *Cancer Cell*, vol. 1, no. 3, pp. 279–288, 2002.

[44] R. Pal, I. Ivanov, A. Datta, M. Bittner, and E. R. Dougherty, "Generating Boolean networks with a prescribed attractor structure," *Bioinformatics*, vol. 21, no. 21, pp. 4021–4025, 2005.

[45] I. Shmulevich, E. R. Dougherty, and W. Zhang, "Control of stationary behavior in probabilistic Boolean networks by means of structural intervention," *Biological Systems*, vol. 10, no. 4, pp. 431–446, 2002.

[46] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Nashua, NH, USA: Athena Scientific, 2001.

[47] C. Derman, *Finite State Markovian Decision Processes*. Orlando, FL, USA: Academic Press, 1970.

[48] E. Altman, *Constrained Markov Decision Processes*. Boca Raton, FL, USA: Chapman and Hall/CRC, 1999.

[49] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[50] A. Datta, A. Choudhary, M. L. Bittner, and E. R. Dougherty, "External control in Markovian genetic regulatory networks," *Machine Learning*, vol. 52, no. 1/2, pp. 169–191, 2003.

[51] T. Akutsu, M. Hayashida, W.-K. Ching, and M. K. Ng, "Control of Boolean networks: Hardness results and algorithms for tree structured networks," *Journal of Theoretical Biology*, vol. 244, no. 4, pp. 670–679, 2007.

[52] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Belmont, MA, USA: Athena Scientific, 1996.

[53] J. N. Tsitsiklis, "Asynchronous stochastic approximation and Q-learning," *Machine Learning*, vol. 16, no. 3, pp. 185–202, 1994.

VITA

Babak Faryabi received the B.S. and M.S. degrees in electrical and computer engineering from the Sharif University of Technology, Tehran, Iran, in 1995 and 1997, respectively. He received his Ph.D. degree from Texas A&M University, College Station, Texas, in August, 2009.

Prior to joining Texas A&M University, he was a research assistant at the University of Toronto, Ontario, Canada, where he was the recipient of the Edward S. Rogers Scholarship from the Department of Electrical and Computer Engineering, University of Toronto, from 2003 to 2005.

Faryabi's research interests include probability theory, statistical methods, signal processing techniques, and their applications to transform our understanding of living organisms as biological systems. He is a member of the IEEE Engineering in Medicine and Biology Society (EMB) and International Society for Computational Biology (ISCB).

Faryabi may be reached at Texas A&M University, Electrical and Computer Engineering Department, College Station, Texas, 77843 – 3128. His email is bfariabi@tamu.edu.