WILEY | Hindawi

Research Article

Optimal Intervention in Semi-Markov-Based Asynchronous Probabilistic Boolean Networks

Qiuli Liu [iD], Qingguo Zeng, Jinghao Huang [iD], and Deliang Li

School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China

Correspondence should be addressed to Qiuli Liu; liuql@m.scnu.edu.cn

Academic Editor: Danilo Comminiello

Synchronous probabilistic Boolean networks (PBNs) and generalized asynchronous PBNs have received significant attention over the past decade as a tool for modeling complex genetic regulatory networks. From a biological perspective, the occurrence of interactions among genes, such as transcription, translation, and degradation, may require a few milliseconds or even up to a few seconds. Such a time delay can be best characterized by generalized asynchronous PBNs. This paper attempts to study an optimal control problem in a generalized asynchronous PBN by employing the theory of average value-at-risk (AVaR) for finite horizon semi-Markov decision processes. Specifically, we first formulate a control model for a generalized asynchronous PBN as an AVaR model for finite horizon semi-Markov decision processes and then solve an optimal control problem for minimizing average value-at-risk criterion over a finite horizon. In order to illustrate the validity of our approach, a numerical example is also displayed.

1. Introduction

Modeling genetic regulatory networks is a core issue in system biology. Numerous models of modeling and understanding genetic regulatory networks have been proposed, which include Boolean networks, Bayesian networks, Petri net, differential equations, computational framework based on the automata and languages theory, and probabilistic Boolean networks (PBNs) [1–8]. Among all the models, Boolean networks and PBNs have gained popularity in modeling real genetic regulatory networks [1, 9–11]. They are also very useful to infer genetic regulatory networks because they can monitor the dynamic behavior in complicated systems based on large amounts of gene expression data. Compared with a Boolean network, which is essentially a deterministic model, a stochastic model is more suitable due to the measurement noise in inferring a gene regulatory network. PBNs are then introduced to cope with the shortcoming. According to the updating time for each gene in the networks, the works on PBNs can be roughly classified into two cases: synchronous

PBNs and asynchronous PBNs. Synchronism means that all genes or nodes simultaneously update their activities. The time parameters of the state processes are discrete, and the time between two successive transitions is equal and fixed. However, synchronous PBNs cannot characterize the real situations. For example, interactions among genes, such as transcription, translation, and degradation, may require a few milliseconds or even up to a few seconds [12]. Motivated by practical considerations, generalized asynchronous PBNs have emerged. In a generalized asynchronous PBN, all the genes update synchronously and the updating period of each gene is randomly selected based on a given distribution. We next briefly describe some existing works on the two cases, respectively.

From different treatment viewpoints, synchronous PBNs and their corresponding intervention methods have been studied in the literatures. For example, due to the fact that the dynamic behavior of the synchronous PBNs can be characterized by the theories of Markov chains and Markov decision processes, the studies of synchronous PBNs have mainly

focused on manipulating external (control) variables that alter the transition probabilities of the network and can be used to help the network avoid undesirable states such as those associated with a disease [9–11]. Specifically, based on the introduction of control inputs, Datta et al. [10] formulate the corresponding control problem as a minimization problem with some costs. Optimal policies are obtained by using the theory of average expected costs and expected discounted cost criteria in infinite horizon Markov decision processes in [11]. Chen and Chen [13] study the control of gene regulatory networks using a Markovian approach, and in the study of Ng et al. [14], the corresponding control model for a gene intervention problem is approximated to drive undesirable states to desirable ones and formulated as a minimization problem with integer variables and continuous variables. To reduce the detrimental side effects of cancer treatment and treatment of patients suffering, Faryabi et al. [15] restrict the expected number of treatments beneath some bounds by using the approach of constrained Markov decision processes; Ching et al. [16] formulate the control problem as a new minimization problem of the distance related to system behaviors with hard constraints in which they add the upper bound for the number of control inputs and compute using a similar approximate method in [17]; Kobayashi and Hiraishi propose an integer programming-based approach for a context-sentive PBN in [18], and they also reduce the optimal control of PBNs to a polynomial optimization problem in [19]. Ivanov and Dougherty consider two kinds of mappings reduction and projection and their effect on the original probability structure of a given PBN in [20]. A review of existing references above shares one common feature: the assumption of synchronous time and fixed observable epoch.

Based on semi-Markov processes, asynchronism in PBNs has been introduced by updating each node based on its period. Faryabi et al. obtain an optimal control policy for semi-Markov asynchronous regulatory networks that minimizes the time that the system spends in undesirable states based on the theory of semi-Markov decision processes in [21]. Liu et al. solve an optimal control problem by choosing optimal constituent Boolean networks in the asynchronous PBN [22]; Liu also describes a control model in an asynchronous PBN as a first passage model in controlled semi-Markov processes [23].

This paper attempts to study a different optimality problem in the framework of generalized asynchronous PBNs. The main motivations of this paper are as follows. First, a fixed and constant sojourn time is assumed in a synchronous PBN, while sojourn times are allowed to follow an arbitrary probability distribution in a generalized asynchronous PBN. Hence, generalized asynchronous PBNs are better able to describe the dynamic behavior of a gene regulatory network because updating the period of each gene is allowed to be made at random points when certain critical events occur, such as a great change in the process of gene expression. Second, random fluctuations in the abundance of molecules in the living cell may lead to noise, thereby affecting its growth and well-being [24–26]. Recent studies suggest that the stochastic

noise in gene expression should be measured by the normalized variance. Due to the limitation of the variance measure, it has been replaced by more complicated risk measures like value-at-risk or AVaR [27–29]. Hence, it is meaningful to control and reduce harmful noise in cellular processes via an AVaR approach. Third, although regulatory intervention has been treated for synchronous PBNs, and the development of intervention policies are based on associated Markov decision processes in [10, 11, 13], therapeutic intervention policies within the context of asynchronous PBNs are few. Hence, the AVaR problem of semi-Markov decision processes in the framework of asynchronous PBNs is unsolved and worth studying.

In this paper, the corresponding generalized asynchronous PBN is described as an AVaR model in finite horizon semi-Markov decision processes with each state variable denoting a GAP at some time. The optimality problem is solved by minimizing AVaR related to system behavior during a finite horizon. In terms of the theory of the AVaR model in finite horizon semi-Markov decision processes [28], the state space is a set of all possible GAPs. The available action space consists of all possible control inputs at a given GAP, and the objective function of the network to be optimized is defined as AVaR, which represents the expected loss upon being within some percentage of the worst-case loss scenario by using a policy starting from an initial state. We give the main ideas of the solution to the optimality problem, which have been formalized in [28], in the following. More precisely, firstly, using a representation of AVaR, we convert the problem of minimizing the AVaR of the finite horizon cost for a semi-Markov decision process into a bilevel optimization problem in which the outer optimization problem is an ordinary problem of minimizing a function of a single variable and the inner optimization problem is to minimize the expected positive deviation of the finite horizon cost for semi-Markov decision processes. According to the results presented in [28], the outer optimization problem can be settled by solving the inner optimization problem. Secondly, we will propose a dynamic programming-based algorithm to solve the inner optimization problem. Thirdly, based on the solution of the inner optimization problem, the existence and computation of an AVaR optimal policy are established by solving the outer optimality problem. Therefore, our optimality problem is actually described as an AVaR model in finite horizon semi-Markov decision processes, and the optimality technique developed well in semi-Markov decision processes [28, 30] can be used to solve it efficiently. The advantage of the method used in this paper is to help biologists in further developing better therapy strategies in synchronous and asynchronous PBNs [31–33].

The structure of this paper is as follows. Section 2 gives a brief introduction of generalized asynchronous PBNs. Section 3 formulates a finite horizon model in semi-Markov decision processes for a generalized asynchronous PBN, and an algorithm for solving the optimization problem is provided in Section 4. A numerical example is shown in Section 5. Finally, a brief summary and comments are put in Section 6.

2. Boolean Networks, Synchronous Probabilistic Boolean Networks, and Generalized Asynchronous Probabilistic Boolean Networks

In this section, we give a brief review on Boolean networks, synchronous probabilistic Boolean networks, and generalized asynchronous probabilistic Boolean networks. For a detailed and complete exposition, the readers are referred to the literatures [4, 31].

A Boolean network $B = (V, F)$ consists of a set of nodes (genes) $V = \{x_1, x_2, \ldots, x_n\}$, $x_i \in \{0, 1\}$, $i = 1, \ldots, n$, and a collection of Boolean functions $F = \{f_1, f_2, \ldots, f_n\}$, where $f_i : \{0, 1\}^n \rightarrow \{0, 1\}$, $i = 1, \ldots, n$. F contains one Boolean function for each gene in V. The function f_i is the predictor function of gene i. Each gene x_i can take on one of two binary values, 0 or 1, corresponding to the case that a particular gene is not expressed or expressed. We denote $x_i(t)$ by the state of gene i at time step t. Thus, the overall state of all the genes in the network at time step t is given by the vector $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))$, also known as the gene activity profile (GAP). The rules of the regulatory interactions among the genes are determined by $x_i(t + 1) = f_i(x(t))$ $(i = 1, 2, \ldots, n)$.

If the genes and the corresponding Boolean functions are given, then a Boolean network is defined. It is not difficult to see that a Boolean network is essentially a deterministic model. To overcome the deterministic rigidity of a Boolean network, the uncertainty of intergene relations are imposed on the dynamic process, which creates synchronous PBNs [14]. Specifically, each gene contains several Boolean functions (predictor functions) $f_i^{(j)} (i = 1, 2, \ldots, l(j))$ to be chosen for determining the next state of the corresponding gene in a synchronous PBN, while it has only one Boolean function in a Boolean network. The probability of choosing the Boolean function $f_i^{(j)}$ is $c_i^{(j)}$, where $c_i^{(j)}$ satisfies the following equation.

$$\sum_{i=1}^{l(j)} c_i^{(j)} = 1 \text{ and } 0 \leq c_i^{(j)} \leq 1 \text{ for } j = 1, 2, \ldots, n. \tag{1}$$

We denote f_i by the ith possible realization, which is of the form $f_i = (f_{i_1}^{(1)}, f_{i_2}^{(2)}, \ldots, f_{i_n}^{(n)})$. Each vector-valued function $f_i = (f_{i_1}^{(1)}, f_{i_2}^{(2)}, \ldots, f_{i_n}^{(n)})$ determines a constituent Boolean network in a synchronous PBN. There are at most $N = \prod_{j=1}^{n} l(j)$ constituent Boolean networks. The dynamics of a synchronous PBN can be represented via a Markov chain. Let α and β be the two column vectors. Then the transition probability satisfies the following equation:

$$P\{\omega(t + 1) = \beta \mid \omega(t) = \alpha\}$$
$$= \sum_{i=1}^{N} P_i \left[\prod_{j=1}^{N} \left(1 - \left| f_{K_{ij}}^{(j)}(x_1, x_2, \ldots, x_n) - x_j \right| \right) \right]. \tag{2}$$

Suppose that α and β take all the possible values in $\{0, 1\}^n$, one can obtain the transition probability matrix.

In a synchronous PBN, all the genes (nodes) simultaneously update when one of the Boolean networks is selected. The time parameters of the state processes are discrete, and the time between two successive transitions is equal and fixed. However, such model is less realistic [26]. From a view of biology, interactions between genes that cause transcription, translation, and degradation may require a few milliseconds or even up to a few seconds [12]. Due to the limitation of the time parameters of the state transitions, the generalized asynchronous PBNs, which are an extension of synchronous PBNs, have gained popularity in modeling genetic regulatory networks [21]. In a generalized asynchronous PBN, all the genes update synchronously and the updating period of each gene is randomly selected based on a given distribution.

To be consistent with aforementioned models, we use the same notation to define generalized asynchronous PBNs. Similar to the synchronous PBN, a generalized asynchronous PBN also contains a set of nodes (genes) $V = \{x_1, x_2, \ldots, x_n\}$, where $x_i \in \{0, 1\}$ and each x_i represents the expression value of a gene, and a collection of vector-valued functions $F = \{f_j\}_{j=1}^N$. If one of the N possible realizations of the generalized asynchronous PBN f_j is chosen, the values of all the genes are simultaneously updated according to the predictors determined by f_j. The gene activity profile is an n-tuple $x(t) = (x_1(t), \ldots, x_n(t))$ giving the values of genes at time t, where $x(t) \in \{0, 1\}^n$. Different from synchronous PBNs, the time parameter of the state process is continuous, the time length between two transitions is random, and the random time follows any distribution in a generalized asynchronous PBN. Such superior properties of the generalized asynchronous PBN can best characterize the real gene regulatory networks. Two consecutive epoch times t_k and t_{k+1} are considered. When the generalized asynchronous PBN occupies the GAP $z(t_k) = x$, a constituent Boolean network is chosen according to some regulatory rule. As a consequence of the constituent Boolean network choice, the network jumps to a new GAP $z(t_{k+1}) = y$ after a sojourn time, whose distribution can be characterized by semi-Markov kernel $Q(x, y, t) = P(y \mid x)H(x, \xi)$. The transition probability satisfies $P(y \mid x) := P(z(t_{k+1}) = y \mid z(t_k) = x)$, ξ is a random variable, and $H(x, \xi)$ denotes its distribution function.

3. A Semi-Markov Decision Process Model for Generalized Asynchronous Probabilistic Boolean Networks

In this section, we will formulate a generalized asynchronous PBN as an AVaR model for finite horizon semi-Markov decision processes with specified definitions of notation as follows.

First, we specify the state of the generalized asynchronous PBN, which is the information to the controller. The state α of the generalized asynchronous PBN at any time $t \geq 0$ denotes the gene activity profile which is a vector $\alpha := (x_1, \cdots, x_n)$, where x_i represents the expression value

of gene i and $x_i \in \{0, 1\}$, $i = 1, 2, \ldots, n$. To facilitate our discussion, we use the following state representation of the network and define

$$x = 1 + \sum_{i=1}^{n} 2^{n-i} x_i, \tag{3}$$

to be the state of the network. Therefore, the corresponding state space is composed of all possible GAPs, that is, $S = \{1, 2, \ldots, 2^n\}$.

Second, a vector consisting of control inputs $\mathbf{u} = (u_1, \ldots, u_m)$ can be regarded as an action available to the controller at state x. The use of such control inputs makes sense from a treatment viewpoint. For instance, in the case of diseases like cancer, auxiliary treatment inputs, such as radiation and chemotherapy, may be applied to move the state vectors away from one which is associated with proliferation of cancer cells. In order to accord with the above decimal number, we take a (the decimal representation of a control input) as an action. We denote $A(x)$ by a set of actions available to the controller when the generalized asynchronous PBN is at state $x \in S$. In other words, $A(x)$ represents the set of all possible treatments, which is assumed to be a finite set $A(x) = \{1, 2, \ldots, 2^m\}$ in this paper.

Third, the generalized asynchronous PBN occupies a state $x \in S$ at some decision epoch and the controller chooses a vector of control inputs $a \in A(x)$ (an action) according to some decision rules. As a consequence of this action choice, the following thing occurs: the generalized asynchronous PBN jumps to a new state y after a sojourn time in x, whose distribution is subject to the semi-Markov kernel. To characterize such event, we need to introduce the semi-Markov decision kernel $Q(t, y \mid x, a)$ as follows. For every $t \in R_+$, x, $y \in S$, and a given action $a \in A(x)$, based on [10, 28], the semi-Markov decision kernel Q is of the form

$$Q(t, y \mid x, a) = H(t \mid x, a) p(y \mid x, a), \tag{4}$$

where $H(t \mid x, a)$ and $p(y \mid x, a)$ denote the distribution functions of the sojourn time and the transition probabilities in mathematics, respectively. In fact, from the definition of the generalized asynchronous PBN and the properties of semi-Markov processes, the generalized asynchronous PBN stays in a given state and a random length of time [28] and its distribution function is denoted by $H(t \mid x, a)$. In the previous section, we have shown that the dynamic behavior of generalized asynchronous PBN could be represented as a semi-Markov process where the state transition was completely specified by all of the possible Boolean functions.

Fourth, the cost function is assumed to be $c(x, a)$, which depends on the current state $x \in S$ and a chosen action $a \in A(x)$.

Fifth, policies are needed to specify decision rules at every jump time. A policy is actually a certain operation rule through which the controller chooses an action. The policies under our optimization problem are based on the decision epochs, as well as the states and actions for the controller. Thus, a policy can be expressed by a function g acting on $R_+ \times S$ with the property that $g(t, x)$ belongs to $A(x)$ for all $(t, x) \in R_+ \times S$. In other words, the function $g(t, x)$ is an action chosen at the current state x and the decision epoch t, which is controlled by the decision-maker using policy g. Let G be the set of all policies.

For requirements, we also introduce some other notation used in this paper. For each $(t, x) \in R_+ \times S$ and $g \in G$, by the well-known Tulcea theorem in [28, 30, 34], there exists a stochastic process $\{T_n, J_n, A_n, n \geq 0\}$, in which T_n, J_n, and A_n denote the nth decision epoch and the state and the action (control input) chosen at the nth decision epoch, respectively.

Under suitable conditions, we define an underlying continuous-time semi-Markov decision process $\{Z(t), U(t), t \in R_+\}$ corresponding to the discrete-time process $\{T_n, J_n, A_n\}$ by

$$Z(t) = J_n, \ U(t) = A_n, \ \text{for } T_n \leq t < T_{n+1}, \ t \in R_+ \text{ and } n \geq 0. \tag{5}$$

Finally, in order to complete our specified model, we need to introduce an optimality criterion. Now, for each $(t, x) \in [0, T] \times S$, we define the value-at-risk (VaR) of finite horizon total cost at level $\gamma \in (0, 1)$ under a policy $g \in G$ by

$$\zeta_\gamma^g(t, x) := \inf \left\{ \lambda \in \mathbb{R} \mid P_{(t,x)}^g \left(\int_t^T c(Z(s), U(s)) ds \leq \lambda \right) \geq \gamma \right\}, \tag{6}$$

and the AVaR of finite horizon total cost at level γ under a policy $g \in G$ by

$$\eta_\gamma^g(t, x) := \frac{1}{1 - \gamma} \int_\gamma^1 \zeta_s^g(t, x) ds. \tag{7}$$

The VaR $\zeta_\gamma^g(t, x)$ can be interpreted as the maximum possible loss for the generalized asynchronous PBN with probability at least γ over a time horizon $[t, T]$. Hence, the AVaR $\eta_\gamma^g(t, x)$ is the expected cost exceeding the VaR $\zeta_\gamma^g(t, x)$.

In this paper, we are interested in minimizing η_γ^g over $g \in G$ and aim to find a policy g^* such that

$$\eta_\gamma^{g*}(t, x) = \eta_\gamma^*(t, x), \tag{8}$$

where $\eta_\gamma^*(t, x) := \inf_{g \in G} \eta_\gamma^g(t, x)$ is the minimum AVaR. Such a policy g^*, when it exists, is called AVaR optimal.

To that end, we have specified the AVaR model for the above generalized asynchronous PBN as follows:

$$\left\{ S, (A(x), x \in S), Q(t, y \mid x, a), c(x, a), \left(\eta_\gamma^g(t, x), g \in G \right) \right\}, \tag{9}$$

where the state space S; the available action set $A(x)$ at state $x \in S$; the semi-Markov kernel $Q(t, y \mid x, a)$ with $x, y \in S$, $t \in$

[0, T], and $a \in A(x)$; and the AVaR $\eta_\gamma^g(t, x)$ have been defined as above.

To solve our optimization problem (8), we need the following alternative representation for AVaR as shown in the literatures [27–29].

Let $g \in G$ and $\gamma \in (0, 1)$. Then, for every $(t, x) \in [0, T] \times S$, it holds that

$$
\eta_\gamma^g(t, x) = \inf_{\lambda \in \mathbb{R}} \left\{ \lambda + \frac{1}{1-\gamma} E_{(t,x)}^g \left[\left[\int_t^T c(Z(s), U(s)) ds - \lambda \right]^+ \right] \right\},
$$

(10)

where the minimum point is attained at $\lambda^*(t, x) = \zeta_\gamma^g(t, x)$. The equality above implies that AVaR is equal to VaR plus the expected losses exceeding VaR divided by the probability of these losses occurring, $1 - \gamma$.

To solve (8), we provide a way of minimizing AVaR. In fact, by (10), we have

$$
\begin{aligned}
\eta_\gamma^*(t, x) &= \inf_{g \in G} \eta_\gamma^g(t, x) \\
&= \inf_{g \in G} \inf_{\lambda \in \mathbb{R}} \left\{ \lambda + \frac{1}{1-\gamma} E_{(t,x)}^g \left[\left[\int_t^T c(Z(s), U(s)) ds - \lambda \right]^+ \right] \right\} \\
&= \inf_{\lambda \in \mathbb{R}} \left\{ \lambda + \frac{1}{1-\gamma} \inf_{g \in G} E_{(t,x)}^g \left[\left[\int_t^T c(Z(s), U(s)) ds - \lambda \right]^+ \right] \right\}.
\end{aligned}
$$

(11)

Based on the theory of [27, 28], to solve the outer optimality problem (i.e., our original optimization problem (8)), we shall first study the inner optimization problem. To this end, we define the expected positive deviation of finite horizon cost from a level λ under a policy $g \in G$ by

$$
J^g(t, x, \lambda) := E_{(t,x,\lambda)}^{\tilde{g}} \left[\int_t^T c(Z(s), U(s)) ds - \lambda \right]^+
$$

$$
\cdot (t, x, \lambda) \in [0, T] \times S \times \mathbb{R}.
$$

(12)

Here, λ can be interpreted as the acceptable cost/loss that the controller expects to spend/lose, and $J^g(t, x, \lambda)$ measures the expected amount by which total cost/loss will exceed an expectation λ over the horizon $[t, T]$. Note that $J^g(t, x, \lambda)$ in (12) depends on the cost level λ. For technical convenience, it is natural to introduce a class of more general policies including the cost level λ. We rewrite the definition of the more general policies denoted by $\tilde{g}$ in lieu of g. Our goal now is, for every fixed $\lambda \in \mathbb{R}$, to minimize $J^{\tilde{g}}(\cdot, \cdot, \lambda)$ over $\tilde{g} \in \tilde{G}$ and to seek an optimal policy $\tilde{g}^* \in \tilde{G}$ (depending on λ) satisfying

$$
J^{\tilde{g}^*}(t, x, \lambda) = J^*(t, x, \lambda) \quad \forall (t, x, \lambda) \in [0, T] \times S \times \mathbb{R},
$$

(13)

where $J^*(t, x, \lambda) = \inf_{\tilde{g} \in \tilde{G}} J^{\tilde{g}}(t, x, \lambda)$ is the value function for the expected positive deviation criterion.

The main goal of this paper is to solve the stochastic optimal control problem above with the help of the theory of semi-Markov decision processes [28, 30]. As described above, using a representation of AVaR in (10), we have converted the problem of minimizing the AVaR of the finite horizon cost for a semi-Markov decision process into a bilevel optimization problem in which the outer optimization problem is an ordinary problem of minimizing a function of a single variable and the inner optimization problem is to minimize the expected positive deviation of the finite horizon cost for semi-Markov decision processes.

4. Solutions to the Optimization Problem

Based on the AVaR model in finite horizon semi-Markov decision processes of the generalized asynchronous PBN described in Section 3, in this section we want to solve the optimality problem and propose a solution method for the optimization problem. We give the main ideas of the solution process as follows. In fact, the algorithm comes from the theory of the AVaR problem for finite horizon semi-Markov decision processes. In order to obtain solutions for the original AVaR criterion (8) in the framework of generalized asynchronous PBNs, we have done it in three steps.

In the first step, we formulate our original problem (8) as a bilevel optimization problem. This step is executed via the representation (10).

In the second step, we solve the inner optimization problem for an arising intermediate criterion (12) in finite horizon semi-Markov decision processes. This treatment is key to analyzing the inner optimization problem.

In the third step, we derive an AVaR optimal policy for the original problem (8) from the solution of the inner optimization problem (12).

From (3.8) in [28], for each $(t, x, \lambda) \in [0, T] \times S \times \mathbb{R}$ and $\tilde{g} \in \tilde{G}$, we rewrite arising intermediate criterion (12) as (14) for the inner optimization problem.

$$
\begin{aligned}
J^{\tilde{g}}(t, x, \lambda) &= E_{(t,x,\lambda)}^{\tilde{g}} \left[\int_t^T c(Z(s), U(s)) ds - \lambda \right]^+ \\
&= \lim_{n \to \infty} E_{(t,x,\lambda)}^{\tilde{g}} \left[\sum_{m=0}^n c(X_m, A_m)((T - T_m)^+ \wedge \Theta_{m+1}) - \lambda \right]^+,
\end{aligned}
$$

(14)

where $\Theta_{m+1} := T_{m+1} - T_m$ denotes the sojourn times between two successive decision epochs.

According to (14), we define $J_{-1}^{\tilde{g}}(t, x, \lambda) := (0 - \lambda)^+ = \lambda^-$ and

$$
J_n^{\tilde{g}}(t, x, \lambda) := E_{(t,x,\lambda)}^{\tilde{g}} \left[\sum_{m=0}^n c(X_m, A_m)((T - T_m)^+ \wedge \Theta_{m+1}) - \lambda \right]^+,
$$

(15)

for every $(t, x, \lambda) \in [0, T] \times S \times \mathbb{R}$ and $n \geq 0$. Clearly, $\lim_{n \to \infty} J_n^{\tilde{g}} = J^{\tilde{g}}$. Hence, we shall calculate $J_n^{\tilde{g}}$ so as to compute $J^{\tilde{g}}$. Moreover, from Lemmas 3.3 and 3.4 in [28], we can obtain various similar properties of $J_n^{\tilde{g}}$ and $J^{\tilde{g}}$.

To analyze the optimality problem, we introduce notations H^a and H as follows: for $(t, x, \lambda) \in [0, T] \times S \times \mathbb{R}$, $a \in A(x)$,

$$
\begin{aligned}
H^a D(t, x, \lambda) &:= (1 - Q(T - t, S \mid x, a))(\lambda - c(x, a)(T - t))^- \\
&\quad + \sum_{y \in S} \int_0^{T-t} Q(ds, y \mid x, a) D(t + s, y, \lambda - c(x, a)s), \\
H D(t, x, \lambda) &:= \min_{a \in A(x)} H^a D(t, x, \lambda).
\end{aligned}
\tag{16}
$$

Next, we are ready to establish the computation and the existence of optimal policies for problem (8). The theorem below develops a value iteration algorithm for calculating the value function J^* and also derives the optimality equation. In addition, with the help of the optimality equation, we show the existence of optimal policies for the problem (13).

Theorem 1. *Under suitable conditions, the following assertions are true.*

(a) *For each $(t, x, \lambda) \in [0, T] \times S \times \mathbb{R}$, let*

$$
\begin{aligned}
D_{-1}^*(t, x, \lambda) &:= \lambda^-, \\
D_{n+1}^*(t, x, \lambda) &:= H D_n^*(t, x, \lambda), \quad n \geq -1.
\end{aligned}
\tag{17}
$$

Then, D_n^* increases in n, and $\lim_{n \to \infty} D_n^* = D^*$.

(b) *D^* is the unique solution in some function space to the optimality equation $D^* = H D^*$.*

(c) *There exists a $\tilde{g} \in \tilde{G}$ such that $D^* = H^{\tilde{g}} D^*$, and such a policy is optimal for the optimality problem.*

Proof. The proof of this theorem is similar to that of Theorem 3.1 and Theorem 3.2 in [28]. For a detailed discussion, consult [28] and the references therein.

Remark 1. We have provided a value iteration algorithm for computing the value function (see Theorem 1 (a)), proven that the value function was a unique solution in some function space to the optimality equation (see Theorem 1 (b)), and shown the existence of optimal policies (see Theorem 1 (c)). This result plays a key role in the derivation of the value iteration algorithm and the existence of optimal policies.

We now return to the original problem (8). Let $\omega(t, x, \lambda) := \lambda + (1/1 - \gamma) u^*(t, x, \lambda)$ and consider the problem

$$
\inf_{\lambda \in \mathbb{R}} \omega(t, x, \lambda) = \inf_{\lambda \in \mathbb{R}} \left[\lambda + \frac{1}{1 - \gamma} u^*(t, x, \lambda) \right].
\tag{18}
$$

The following theorem relates optimal policies of the inner optimization problem to AVaR optimal policies of the original problem.

Theorem 2 (see Theorem 3.4 in [28]). *Under suitable conditions, there exists a solution λ^* (depending on (t, x)) of problem (18), and the policy $g^*(\cdot, \cdot) := \tilde{g}^*(\cdot, \cdot, \lambda^*(\cdot, \cdot)) \in G$ is AVaR optimal for the original problem, where $\tilde{g}^* \in \tilde{G}$ is an optimal policy for the optimality problem (13).*

Remark 2. Based on the solution of the inner optimization problem, the existence and computation of an AVaR optimal policy in a generalized asynchronous PBN are established in Theorem 2.

Next, we will propose a value iteration algorithm for computing the optimal value functions based on Theorems 1 and 2 above. Moreover, optimal policies for the optimization problem are constructed in the course of the algorithm. According to [28], the details and the derivation are shown as follows.

Step I. Given an $\epsilon > 0$, and set $n = 0$. Let $D_0(t, x, \lambda) := \lambda^-$ for each $(t, x, \lambda) \in [0, T] \times S \times \mathbb{R}$.

Step II. For each $(t, x, \lambda) \in [0, T] \times S \times \mathbb{R}$, compute $D_{n+1}(t, x, \lambda)$ by

$$
\begin{aligned}
D_{n+1}(t, x, \lambda) = \min_{a \in A(x)} \Bigg[&\left(1 - \sum_{y \in S} Q(T - t, y \mid x, a) \right) \\
&\cdot (\lambda - c(x, a)(T - t))^- \\
&+ \sum_{y \in S} \int_0^{T-t} Q(ds, y \mid x, a) D_n \\
&\cdot (t + s, y, \lambda - c(x, a)s) \Bigg].
\end{aligned}
\tag{19}
$$

Step III. If $D_{n+1} - D_n < \epsilon$, go to Step IV. Otherwise, increment n by 1 and return to Step II.

Step IV. For each $(t, x) \in [0, T] \times S$, find the minimum point $\lambda^*(t, x)$ of the function $\lambda \to \lambda + (1/1 - \gamma) D_{n+1}(t, x, \lambda)$, and stop.

5. Numerical Implementation

In this section, a numerical example on the generalized asynchronous PBN is conducted to illustrate our results. From the results presented in [28], the outer optimization problem can be settled by solving the inner optimization problem. We first solve the inner optimization problem using a dynamic programming-based algorithm in Section 4. Then, based on the solution of the inner optimization problem, the existence and computation of an AVaR optimal policy are established by solving the outer optimization problem.

<div style="text-align:center">TABLE 1: Truth table.</div>

$x_1x_2x_3$	$f_1^{(1)}$	$f_2^{(1)}$	$f_1^{(2)}$	$f_2^{(2)}$	$f_1^{(3)}$	$f_2^{(3)}$
0 0 0	0	0	1	1	1	0
0 0 1	0	0	1	0	1	0
0 1 0	0	0	1	0	0	1
0 1 1	0	0	0	0	0	1
1 0 0	1	1	1	0	1	0
1 0 1	1	1	1	1	1	0
1 1 0	1	0	1	1	1	0
1 1 1	1	1	1	1	0	1
$c_j^{(i)}$	0.5	0.5	0.5	0.5	0.4	0.6

<div style="text-align:center">TABLE 2: Truth table.</div>

u	a	x_1x_2	x	$f_{(1,u)}^{(1)}$	$f_{(2,u)}^{(1)}$	$f_{(1,u)}^{(2)}$	$f_{(2,u)}^{(2)}$
0	1	0 0	1	1	1	1	0
0	1	0 1	2	1	0	1	0
0	1	1 0	3	1	0	0	1
0	1	1 1	4	0	0	0	1
1	2	0 0	1	1	0	1	0
1	2	0 1	2	1	1	1	0
1	2	1 0	3	1	1	1	0
1	2	1 1	4	1	1	0	1
$c_j^{(i)}$				0.5	0.5	0.4	0.6

Consider a generalized asynchronous PBN with three genes, x_1, x_2, and x_3. There are two Boolean functions $f_1^{(i)}$ and $f_2^{(i)}$ associated with each gene i. These boolean functions are provided by the truth table in Table 1. From this table, we have eight possible Boolean networks.

Suppose that x_1 is a control input, taking the value 0 or 1. For the new generalized asynchronous PBN, to be consistent with the notation introduced in Section 2, the variables x_1, x_2, and x_3 are renamed as follows: gene x_1 now becomes u while genes x_2 and x_3 become x_1 and x_2, respectively. The information of Boolean functions can be found in Table 2, which shows that there are four possible Boolean networks.

From Table 2, we know that $n = 2$ in this example. Therefore, the state space can be assumed as $S = \{1,2,3,4\}$, and the set of the admissible actions is given by $A(x) = \{1, 2\}$, $x \in S$.

The semi-Markov kernel Q is of the form $Q(t, y \mid x, a) = H(t \mid x, a)p(y \mid x, a)$ for every $t \in R_+$, x, $y \in S$, and $a \in A(x)$. Suppose that the corresponding distribution functions of the sojourn time $H(t \mid x, a)$ are given by

$$
\begin{aligned}
&H(t \mid 1, 1) = \begin{cases} \dfrac{t}{15}, & t \in [0, 15], \\ 1, & t > 15; \end{cases} \\
&H(t \mid 1, 2) = \begin{cases} \dfrac{t}{15.2}, & t \in [0,15.2], \\ 1, & t > 15.2, \end{cases} \\
&H(t \mid 2, 1) = 1 - e^{-0.1t}, \quad t \in R_+; \\
&H(t \mid 2, 2) = 1 - e^{-0.7t}, \quad t \in R_+, \\
&H(t \mid 3, 1) = \begin{cases} \dfrac{t}{60}, & t \in [0, 60], \\ 1, & t > 60; \end{cases} \\
&H(t \mid 3, 2) = \begin{cases} \dfrac{t}{20}, & t \in [0, 20], \\ 1, & t > 20. \end{cases} \\
&H(t \mid 4, 1) = 1 - e^{-0.15t}, \quad t \in R_+; \\
&H(t \mid 4, 2) = 1 - e^{-0.4t}, \quad t \in R_+;
\end{aligned}
\tag{20}
$$

As described in Section 2, we can derive transition probabilities $p(y \mid x, a)$ based on the truth Table 2 and the involved probabilistic parameters.

Moreover, assume that the cost rates are given by

$$
\begin{aligned}
c(1, 1) &= \bar{r}(1) - \bar{c}(1) = 9 - 1 = 8, \\
c(1, 2) &= \bar{r}(1) - \bar{c}(2) = 9 - 2 = 7, \\
c(2, 1) &= \bar{r}(2) - \bar{c}(1) = 4 - 1 = 3, \\
c(2, 2) &= \bar{r}(2) - \bar{c}(2) = 4 - 2 = 2, \\
c(3, 1) &= \bar{r}(3) - \bar{c}(1) = 5 - 1 = 4, \\
c(3, 2) &= \bar{r}(3) - \bar{c}(2) = 5 - 2 = 3, \\
c(4, 1) &= \bar{r}(4) - \bar{c}(1) = 7 - 1 = 6, \\
c(4, 2) &= \bar{r}(4) - \bar{c}(2) = 7 - 2 = 5.
\end{aligned}
\tag{21}
$$

Note that the distribution functions and the cost rates are completely arbitrary. In a real-world example, this information would be obtained from biologists.

For this generalized asynchronous PBN, we want to minimize the AVaR over a time horizon $[0, T]$ with respect to a given confidence level γ.

In addition, we set $\epsilon = 10^{-3}$, $T = 15$, and $\gamma = 0.95$ and discretize the time interval $[0,15]$ and the cost level interval $[0,150]$. Then, we implement the algorithm stated in Section 4 and obtain some datas on the functions $H^a D^*$ (see Figures 1–4). To be specific, we analyze the datas on $H^a D^*(0, x, \lambda)$, $H^a D^*(2.5, x, \lambda)$, $H^a D^*(5, x, \lambda)$, and $H^a D^*(10, x, \lambda)$ as examples, which are shown in Figures 1–4 below.

Comparing the data on $H^a D^*(t, x, \lambda)$ under admissible action a for every $(t, x, \lambda) \in [0, 15] \times \{1,2,3,4\} \times [0,150]$, one may obtain an optimal policy $\tilde{g}^*$ for the inner optimization problem. For example, in light of figures, we can define $\tilde{g}^*$ by

$$
\tilde{g}^*(0, 1, \lambda) = \begin{cases} 2, & \lambda \le 62.9, \\ 1, & \lambda > 62.9, \end{cases}
$$

$$
\tilde{g}^*(0, 2, \lambda) = \begin{cases} 2, & \lambda \le 20.1, \\ 1, & \lambda > 20.1, \end{cases}
$$
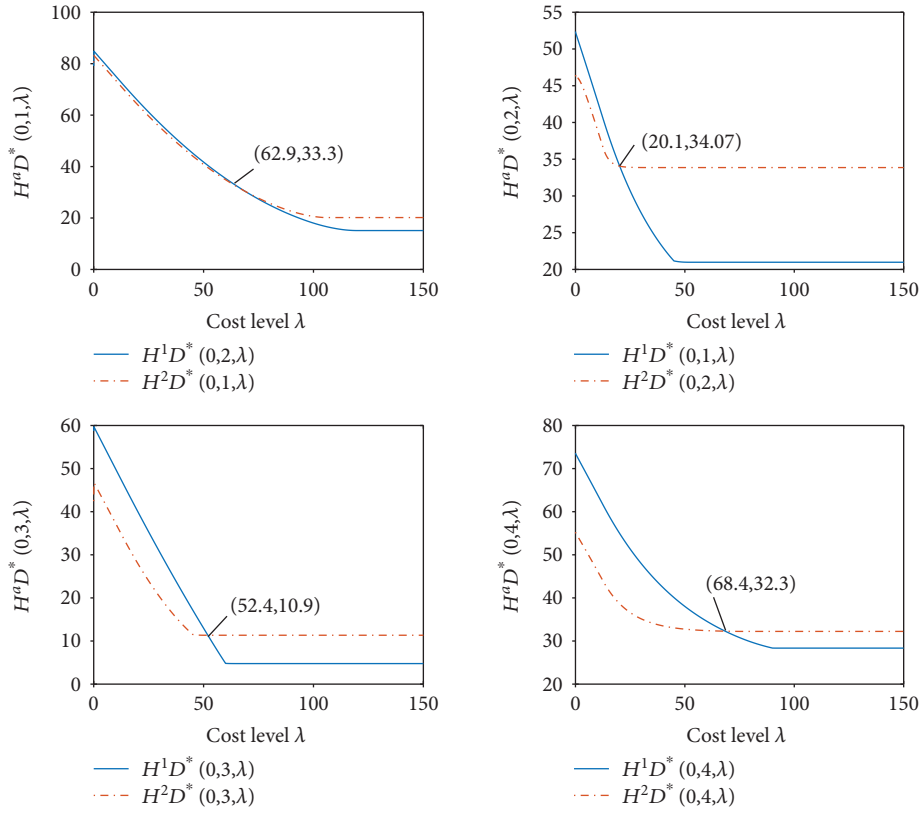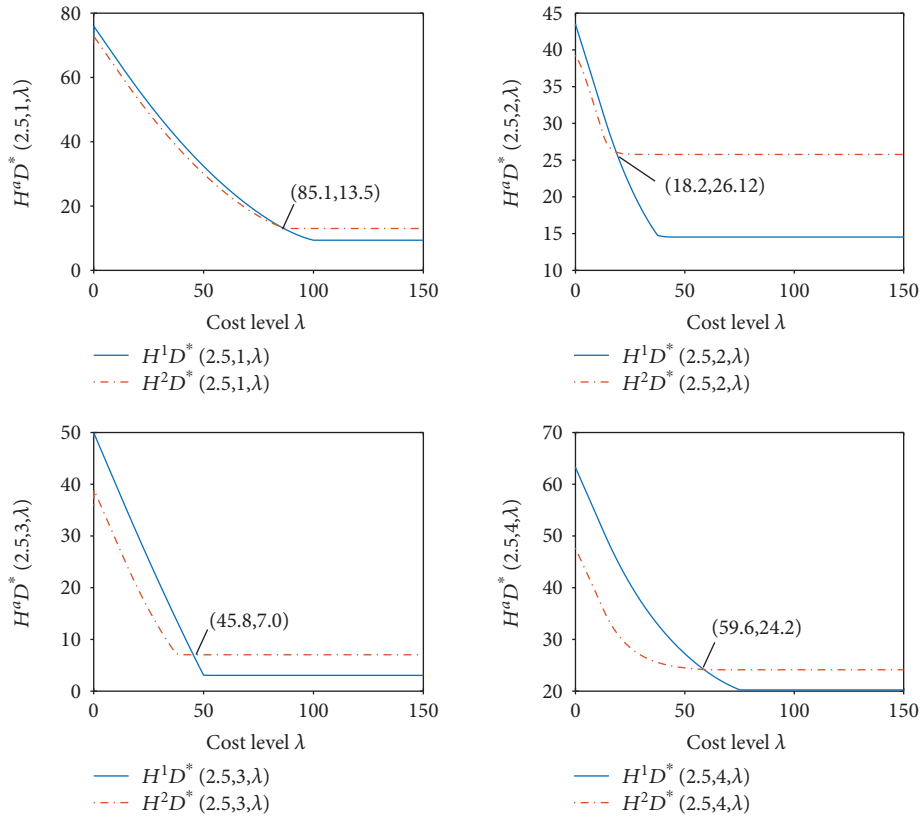
FIGURE 1: The function $H^a D^*(0, x, \lambda)$.

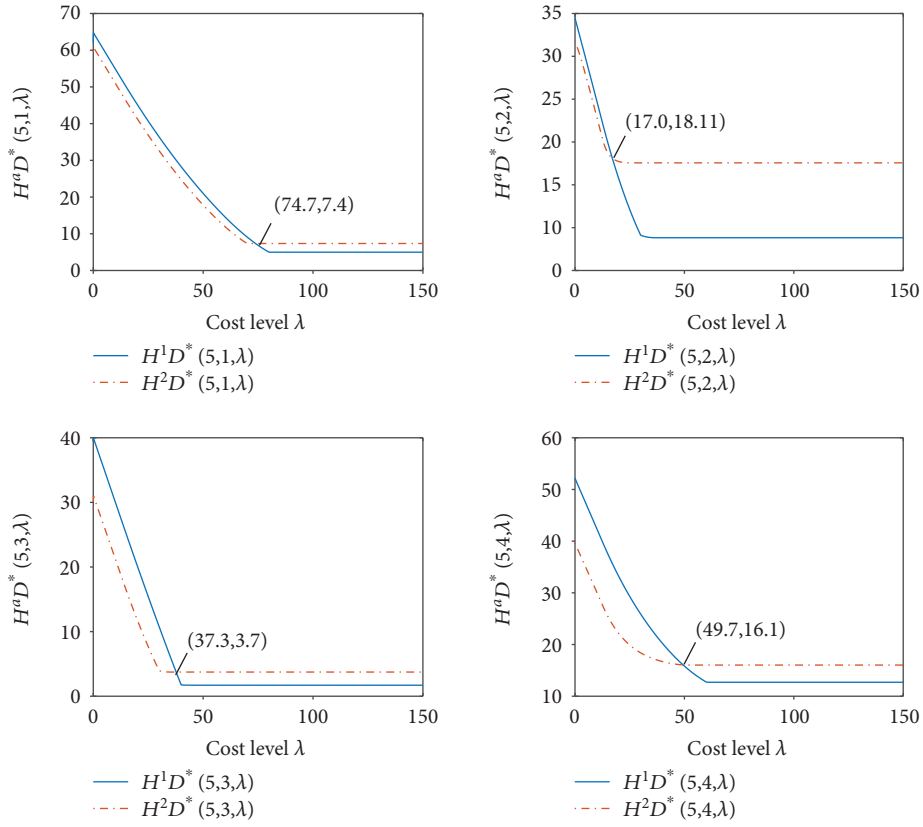FIGURE 2: The function $H^a D^*(2.5, x, \lambda)$.
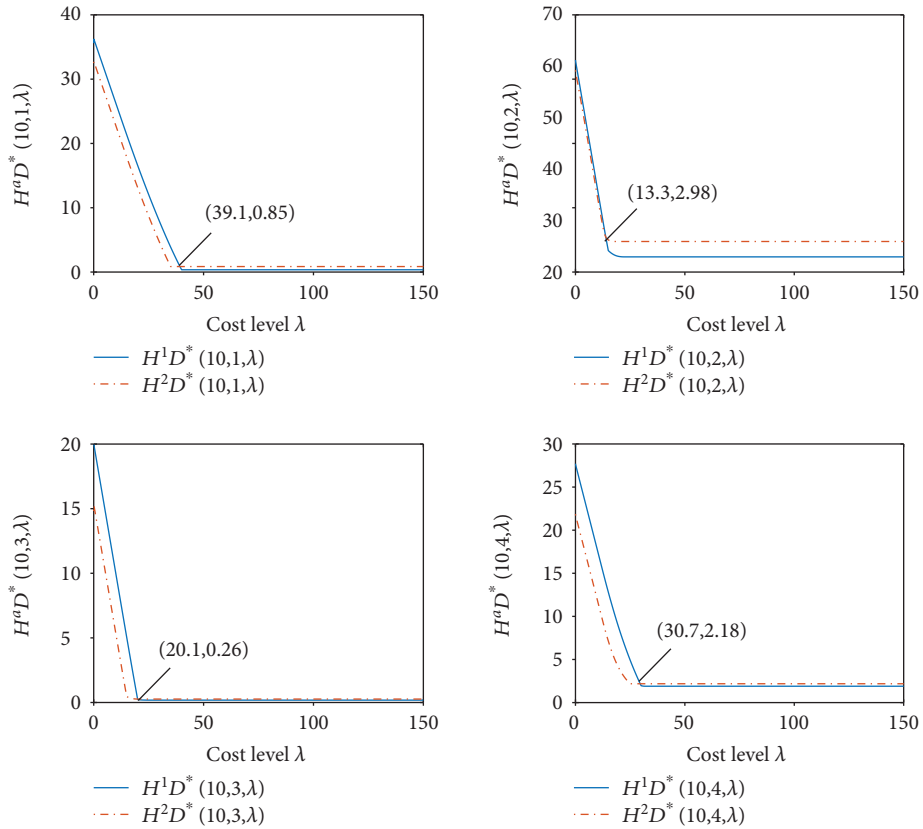
Figure 3: The function $H^a D^*(5, x, \lambda)$.
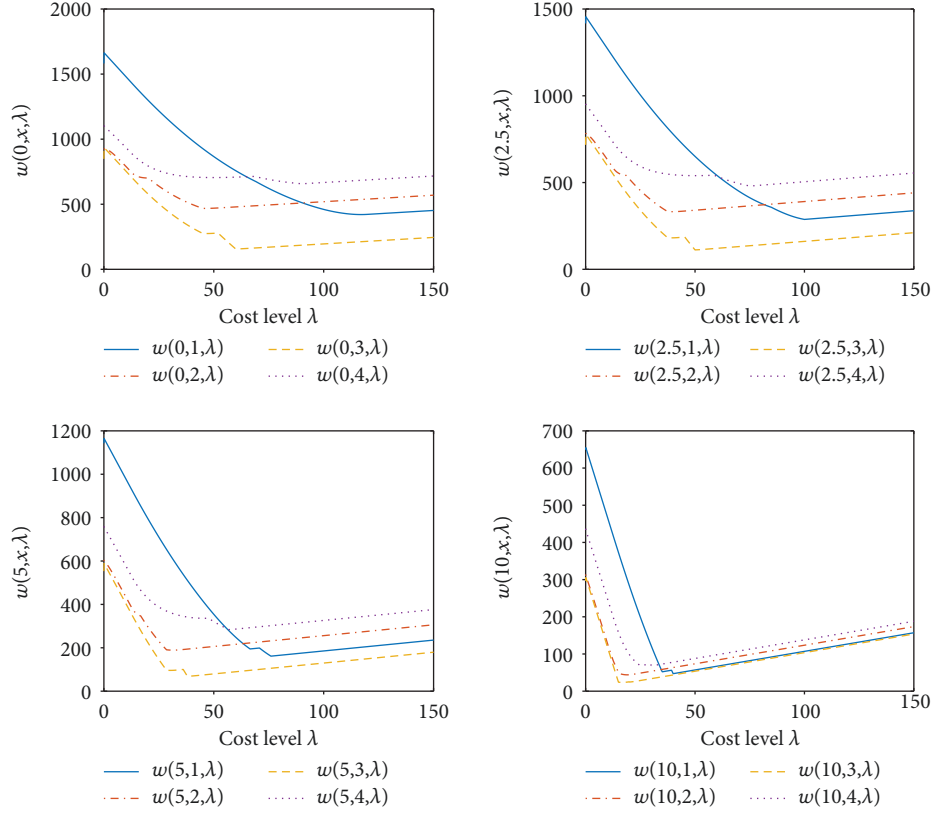

Figure 4: The function $H^a D^*(10, x, \lambda)$.

Figure 5: The function $\omega(t, x, \lambda)$.

$$\tilde{g}^*(0, 3, \lambda) = \begin{cases} 2, & \lambda \leq 52.4, \\ 1, & \lambda > 52.4, \end{cases}$$

$$\tilde{g}^*(0, 4, \lambda) = \begin{cases} 2, & \lambda \leq 68.4, \\ 1, & \lambda > 68.4, \end{cases}$$

$$\tilde{g}^*(2.5, 1, \lambda) = \begin{cases} 2, & \lambda \leq 85.1, \\ 1, & \lambda > 85.1, \end{cases}$$

$$\tilde{g}^*(2.5, 2, \lambda) = \begin{cases} 2, & \lambda \leq 18.5, \\ 1, & \lambda > 18.5, \end{cases}$$

$$\tilde{g}^*(2.5, 3, \lambda) = \begin{cases} 2, & \lambda \leq 45.8, \\ 1, & \lambda > 45.8, \end{cases}$$

$$\tilde{g}^*(2.5, 4, \lambda) = \begin{cases} 2, & \lambda \leq 59.6, \\ 1, & \lambda > 59.6, \end{cases}$$

$$\tilde{g}^*(5, 1, \lambda) = \begin{cases} 2, & \lambda \leq 74.7, \\ 1, & \lambda > 74.7, \end{cases}$$

$$\tilde{g}^*(5, 2, \lambda) = \begin{cases} 2, & \lambda \leq 17, \\ 1, & \lambda > 17, \end{cases}$$

$$\tilde{g}^*(5, 3, \lambda) = \begin{cases} 2, & \lambda \leq 37.3, \\ 1, & \lambda > 37.3, \end{cases}$$

$$\tilde{g}^*(5, 4, \lambda) = \begin{cases} 2, & \lambda \leq 49.7, \\ 1, & \lambda > 49.7, \end{cases}$$

$$\tilde{g}^*(10, 1, \lambda) = \begin{cases} 2, & \lambda \leq 39.1, \\ 1, & \lambda > 39.1, \end{cases}$$

$$\tilde{g}^*(10, 2, \lambda) = \begin{cases} 2, & \lambda \leq 13.2, \\ 1, & \lambda > 13.2, \end{cases}$$

$$\tilde{g}^*(10, 3, \lambda) = \begin{cases} 2, & \lambda \leq 20.1, \\ 1, & \lambda > 20.1, \end{cases}$$

$$\tilde{g}^*(10, 4, \lambda) = \begin{cases} 2, & \lambda \leq 30.7, \\ 1, & \lambda > 30.7. \end{cases}$$

$$(22)$$

According to the definition of policy, $\tilde{g}^*$ is an optimal policy for the inner optimization problem.

Moreover, we find the minimum point $\lambda^*(t, x)$ of the function $\lambda \mapsto \omega(t, x, \lambda)$ with $\gamma = 0.95$. Figure 5 above gives the graphs of $\omega(t, x, \lambda)$ with $t = 0, 2.5, 5, 10$. From Figure 5,

it is easy to see the minimum points $\lambda^*(t, x)$ with $t = 0, 2.5, 5, 10$ and $x = 1, 2, 3, 4$, that is,

$$\lambda^*(0, 1) = 117.4,$$
$$\lambda^*(0, 2) = 45,$$
$$\lambda^*(0, 3) = 60,$$
$$\lambda^*(0, 4) = 90,$$
$$\lambda^*(2.5,1) = 100,$$
$$\lambda^*(2.5,2) = 37.5,$$
$$\lambda^*(2.5,3) = 50,$$
$$\lambda^*(2.5,4) = 75,$$
$$\lambda^*(5, 1) = 76,$$
$$\lambda^*(5, 2) = 28.4,$$
$$\lambda^*(5, 3) = 38,$$
$$\lambda^*(2.5,4) = 75,$$
$$\lambda^*(10, 1) = 40,$$
$$\lambda^*(10, 2) = 15.9,$$
$$\lambda^*(10, 3) = 20.1,$$
$$\lambda^*(10, 4) = 30.$$

(23)

For other $t \in [0, 15]$, the minimum points $\lambda^*(t, x)$ can be similarly calculated. Then, the policy $g^*(t, x) := \tilde{g}^*(t, x, \lambda^*(t, x)) \in G$ is AVaR optimal. For example,

$$g^*(t, x) = 2 \text{ for all } (t, x) \in [0, 15] \times \{1,2,3,4\}. \quad (24)$$

Moreover, we have also computed AVaR $\eta_{0.95}^*(t, x)$ and VaR $\zeta_{0.95}^{g*}(t, x)$ at level 0.95 as shown in Tables 3 and 4.

When the AVaR optimal policy g^* is applied, we will have minimum AVaR $\eta_{0.95}^*(t, x)$ at level 0.95, which represents the minimal expected value of the worst 5% costs of the generalized asynchronous PBN over $[t, T]$. For example, if the network starts from state 1, the minimum mean value of the worst 5% costs over $[0, 15]$ is 420.9. On the other hand, Tables 3 and 4 indicate that the computed VaR is less than the computed AVaR in all scenarios we considered. When g^* is applied, for example, starting from state 1, the VaR of costs of the generalized asynchronous PBN over [0,15] at level 0.95 is 117.4, which represents the maximum costs of the generalized asynchronous PBN over horizon [0,15] with a confidence level 95%. Note that, in minimizing AVaR for this example, we have simultaneously improved VaR. However, it is not clear if the obtained VaR $\zeta_{0.95}^{g*}(t, x)$ is the minimum VaR.

6. Conclusion and Discussion

In this paper, we have proposed a systematic approach to design optimal intervention policies by considering some generalized asynchronous PBNs. Such an approach can

TABLE 3: Some data of AVaR $\eta_{0.95}^*(t, x)$.

t x	0	2.5	5	10
1	420.9	287.6	218.3	66.4
2	553.6	408.0	249.2	68.8
3	181.3	132.0	82.5	29.2
4	761.0	576.0	364.0	106.2

TABLE 4: Some data of VaR $\zeta_{0.95}^{g*}(t, x)$.

t x	0	2.5	5	10
1	117.4	100	76	40
2	45	37.5	28.4	15.9
3	60	50	38	0.1
4	90	75	57.1	30

characterize the time delay among gene interactions. Our analytic results can be analyzed automatically or by medical professionals in order to develop better strategies to improve the quality of medical activities.

It is conceivable that results from this paper may be integrated into identifying potential drugs and treatment methods to achieve the best curative effect on the treatment cancer in the future. Although all the facts indicate that the mechanism of the related applications which is not clear now remain to be developed, the proposed method is only a useful attempt but is useful to explore a new and deeper understanding of biological systems.

Improvement on the proposed method includes further investigation in the following three main areas: computational complexity, classification of states, and the distribution functions of the sojourn time. When the number of states and time points become large, the computational complexity of the algorithm increases. When the number of nodes n and the time horizon T are fixed, the computational complexity of the algorithm is $O(2^n)$. The more discussions about the computational complexity of the algorithm can consult [28] and the references therein. As the lack of data has prohibited the inference of any realistic alternative asynchronous models, we have assumed that both the classification of states, the distribution functions of the sojourn time, and the cost function are known. Hence, improving computation efficiency of the algorithm for obtaining the optimal policy for the formulated control model and theoretically getting a reasonable classification of states and the distribution functions of the sojourn time are left as one of our future works.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] 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.

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

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

[4] 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.

[5] U. Alon, *An Introduction to Systems Biology*, Chapman and Hall, 2007.

[6] L. J. Steggles, R. Banks, and A. Wipat, "Modelling and analysing genetic networks: from Boolean networks to Petri nets," in *Computational Methods in Systems Biology*, vol. 4210 of Lecture Notes in Computer Science, , pp. 127–141, Springer, 2006.

[7] P. C. Y. Chen, "A computational model of a class of gene networks with positive and negative controls," *Biosystems*, vol. 73, no. 1, pp. 13–24, 2004.

[8] P. C. Y. Chen, "A discrete-event approach to transcription control with dynamic event-controllability," *Journal of Computational Biology*, vol. 13, no. 9, pp. 1532–1545, 2006.

[9] A. Datta, R. Pal, and E. Dougherty, "Intervention in probabilistic gene regulatory networks," *Current Bioinformatics*, vol. 1, no. 2, pp. 167–184, 2006.

[10] 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.

[11] 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.

[12] F. Greil, B. Drossel, and J. Sattler, "Critical Kauffman networks under deterministic asynchronous update," *New Journal of Physics*, vol. 9, no. 10, 2007.

[13] P. C. Y. Chen and J. W. Chen, "A Markovian approach to the control of genetic regulatory networks," *Biosystems*, vol. 90, no. 2, pp. 535–545, 2007.

[14] M. K. Ng, S. Q. Zhang, W. K. Ching, and T. Akutsu, "A control model for Markovian genetic regulatory networks," in *Transactions on Computational Systems Biology V*, pp. 36–48, Springer, 2006.

[15] B. Faryabi, G. Vahedi, J. F. Chamberland, A. Datta, and E. R. Dougherty, "Optimal constrained stationary intervention in gene regulatory networks," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2008, 10 pages, 2008.

[16] W. K. Ching, S. Q. Zhang, Y. Jiao, T. Akutsu, N. K. Tsing, and A. S. Wong, "Optimal control policy for probabilistic Boolean networks with hard constraints," *IET Systems Biology*, vol. 3, no. 2, pp. 90–99, 2009.

[17] W.-K. Ching, S. Zhang, M. K. Ng, and T. Akutsu, "An approximation method for solving the steady-state probability distribution of probabilistic Boolean networks," *Bioinformatics*, vol. 23, no. 12, pp. 1511–1518, 2007.

[18] K. Kobayashi and K. Hiraishi, "An integer programming approach to optimal control problems in context-sensitive probabilistic Boolean networks," *Automatica*, vol. 47, no. 6, pp. 1260–1264, 2011.

[19] K. Kobayashi and K. Hiraishi, "Optimal control of probabilistic Boolean networks using polynomial optimization," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95.A, no. 9, pp. 1512–1517, 2012.

[20] I. Ivanov and E. R. Dougherty, "Reduction mappings between probabilistic Boolean networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 1, 7 pages, 2004.

[21] B. Faryabi, J. F. Ç. 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.

[22] Q. Liu, X. Guo, and T. Zhou, "Optimal control for generalized asynchronous probabilistic Boolean networks," *Current Bioinformatics*, vol. 7, no. 1, pp. 56–62, 2012.

[23] Q. Liu, "Optimal control for gene regulatory networks based on controlled semi-Markov process," *Current Bioinformatics*, vol. 9, no. 4, pp. 380–388, 2014.

[24] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins, "Stochasticity in gene expression: from theories to phenotypes," *Nature Reviews Genetics*, vol. 6, no. 6, pp. 451–464, 2005.

[25] J. Paulsson, "Models of stochastic gene expression," *Physics of Life Reviews*, vol. 2, no. 2, pp. 157–175, 2005.

[26] I. Harvey and B. Terry, "Time out of joint: attractors in asynchronous random boolean networks," in *Proceedings of the Fourth European Conference on Artificial Life*, pp. 67–75, Brighton, UK, 1997, MIT Press.

[27] N. Bauerle and J. Ott, "Markov decision processes with average-value-at-risk criteria," *Mathematical Methods of Operations Research*, vol. 74, no. 3, pp. 361–379, 2011.

[28] Y. Huang and X. Guo, "Minimum average value-at-risk for finite horizon semi-Markov decision processes in continuous time," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 1–28, 2016.

[29] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[30] N. Limnios and G. Oprişan, *Semi-Markov Processes and Reliability*, Springer Science and Business Media, 2012.

[31] 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.

[32] E. R. Dougherty and A. Datta, "Genomic signal processing: diagnosis and therapy," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 107–112, 2005.

[33] G. Vahedi, B. Faryabi, J. F. Chamberland, A. Datta, and E. R. Dougherty, "Optimal intervention strategies for cyclic therapeutic methods," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 2, pp. 281–291, 2009.

[34] X. Guo and O. Hernández-Lerma, *Continuous-Time Markov Decision Processes*, Springer, Berlin, Heidelberg, 2009.