Published in IET Systems Biology
Received on 1st May 2008
Revised on 22nd September 2008
doi: 10.1049/iet-syb.2008.0121

ISSN 1751-8849

Robust control of uncertain context-sensitive probabilistic Boolean networks

S.Z. Denic[1] *B. Vasic*[1] *C.D. Charalambous*[2,*] *R. Palanivelu*[3]

[1]*Department of Electrical and Computer Engineering, University of Arizona, Tucson, USA*
[2]*Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus*
[3]*Department of Plant Sciences, University of Arizona, Tucson, USA*
[*]*SITE, University of Ottawa, Ottawa, Canada*
E-mail: sdenic@ece.arizona.edu

Abstract: Uncertainty is an intrinsic phenomenon in control of gene regulatory networks (GRNs). The presence of uncertainty is related to impreciseness of GRN models due to: (1) Errors caused by imperfection of measurement devices and (2) Models' inability to fully capture a complex structure of the GRN. Consequently, there is a discrepancy between actual behaviour of the GRN and what is predicted by its mathematical model. This can result in false control signals, which can drive a cell to an undesirable state. To address the problem of control under uncertainties, a risk-sensitive control paradigm is proposed. Robustness is accomplished by minimisation of the mean exponential cost as opposed to, for instance, minimisation of the mean square cost by risk-neutral controllers. The authors derive an optimal risk-sensitive controller when a GRN is modelled by a context-sensitive probabilistic Boolean network (CSPBN). By using a relation between the relative entropy and free-energy, a relative stability of the cost achieved by the risk-sensitive controller is demonstrated when the distribution of the CSPBN attractors is perturbed, as opposed to the cost of the risk-neutral controller that exhibits increase. The use of the relation between the relative entropy and free-energy to analyse the influence of a particular attractor on the robustness of the controller is studied. The efficiency of the risk-sensitive controller is tested for the CSPBN obtained from the study of malignant melanoma.

1 Introduction

External control or intervention (in terminology of [1]) of gene regulatory networks (GRNs) is one of the major research topics in systems biology [2]. External control is accomplished by introducing exogenous control variables that change the internal dynamics of the cell, for example, by turning off targeted genes in the so-called anti-sense therapy. Without external control, a GRN follows its own internal dynamics in accomplishing a number of complex tasks such as a cell growth, division and reproduction. Although the self-regulation structure of a GRN is very robust [3], it may happen that some internal (for instance, spontaneous mutation) or external (for instance, damage caused by toxic chemical or radiation, so-called induced mutation) event changes its dynamics. These events may bring cells to some undesirable, abnormal states. The role of external GRN control is to prevent or to reverse the undesirable state of the cell into a normal one. Applications of the GRN control include a design of optimal therapies and interventions for controlling genetic related diseases or improving crop productivity [2].

The way a GRN is externally controlled depends on its mathematical model. Here, we concentrate on the control of GRNs described by probabilistic Boolean networks (PBNs) [4]. A PBN is a dynamical finite state discrete-time model of a GRN, which defines how the activity of each gene (gene expression) evolves in time as a function of the activity of other genes in the GRN. It represents a stochastic generalisation of a deterministic Boolean network (BN) model [5]. In essence, a PBN can be understood as a collection of deterministic BNs where: (1) At each time instant, the activity of each gene is determined by one BN from the collection, (2) At the next time instant, another BN is chosen according to certain probability distribution

from the collection, which governs gene activities at that time instant. An extension of the PBN is called a context-sensitive PBN (CSPBN). It differs from the PBN for two reasons: (1) Each gene is allowed to change its activity with a small probability p at each time instant regardless of which constituent BN is active at the moment, and (2) Switching between constituent BNs occurs with a small probability q such that a PBN may remain in the same constituent BN for a longer time interval. Implications of these two extensions shall be discussed later.

To date, several approaches have been proposed for the control of PBNs. In [6], if the PBN occupies some undesirable state, a control is achieved by influencing the activity of targeted genes that brings the PBN into another transient state from which the network evolves to a desirable stable state. In [7], a more sophisticated control is proposed that alters functional dependencies between genes causing changes in the steady-state distribution of the PBN. In practice, this can be achieved by injecting certain chemical agents into the cell. Previous strategies have limited effect as they represent transient interventions. The problem of PBN control becomes better defined by realising that PBN is in essence a finite state Markov chain [4]. Therefore well-known optimal control techniques may be good candidates for the PBN control. In [8, 9], the PBN control is formulated and solved in terms of the classical dynamic programming algorithm over a finite horizon. Further consideration of the finite horizon control is addressed in [10], under the constraint on the maximum number of controls that can be exercised. For the control of CSPBN see [11, 12].

The underlying assumptions of the previous work are the following: (1) The probability distribution of a controlled PBN is completely known and (2) The PBN accurately represents the behaviour of a corresponding GRN. However, as recognised in [1], this is not true in a sense that there is always a discrepancy between a PBN model and actual GRN. The discrepancy is due to two main reasons, measurement errors and simplicity of a model. Measurement errors occur because of the inherent limitations of the measurement devices such as oligonucleotide arrays or cDNA microarrays. In the case of cDNA microarrays, different types of noises impede the quality of measurements (for an excellent treatment see [13]). In addition, very often measurements are performed *in vitro*, which may skew a real information and lead to wrong conclusions [14]. Modelling is another reason for uncertainty. Here, the uncertainty arises because, in practice, only relatively simple GRN models can be employed, tractable from the computational point of view. (Since a state space of even moderate-in-size GRN can be very large, more sophisticated models would require computers with high computational power.) This model simplification inevitably introduces a mismatch between a model and reality, leading to uncertainty.

The presence of uncertainty may cause a poor efficiency of external GRN control. If non-robust control techniques are applied in the presence of uncertainty, a severe degradation of the performance may occur comparing to the performance when the GRN is completely known. In general, the larger the distance between the nominal model (a model obtained from limited measurements or belief) and the true model (real behaviour of the GRN), the poorer the efficiency of the non-robust control system. When a GRN is modelled by a PBN, the control under the uncertainty depends on the distance between the nominal PBN distribution and the true distribution of the GRN.

Previous discussion suggests that an optimal robust control should accomplish a control goal despite a discrepancy between the nominal and true distribution of the PBN. To achieve the robustness in the control of GRNs, we propose a well-known risk-sensitive control approach [15, 16] (for other approaches dealing with uncertainties in PBNs see [17]). The application of a risk-sensitive control for GRNs is supported by the fact that a CSPBN is a special case of a Markov chain.

A risk-sensitive controller minimises the average exponential cost as opposed to a standard risk-neutral controller that minimises, for instance, a mean square cost. An advantage of the risk-sensitive control is that when the exponential cost is minimised, then in addition to the usual cost (mean square cost), all higher moments of the cost are minimised as well. It can be proved that minimisation of the average exponential cost is equivalent to the H^{∞} robust control approach (for further details see [18]) explaining the robustness of the risk-sensitive controllers. A virtue of the risk-sensitive controller is that the value of the achieved cost remains relatively stable even when the distance between the nominal and true distribution of the PBN increases, which is not the case for risk-neutral controllers.

Following an approach found in [18] (see also [19]), we state the optimal control problem of partially observed (observed in noise) CSPBNs and thereafter derive the optimal risk-sensitive controller. Since we consider a finite horizon case, a dynamical programming gives the optimal controller equation. The solution is based on the discrete version of Girsanov's theorem and the change-of-measure technique.

The performance of the risk-sensitive controller is estimated by using a relation between the free-energy and the relative entropy [20], which is the main contribution of the paper. The relative entropy is a natural measure to assess the performance of a controller subject to uncertainties since it represents a measure of distance between two absolutely continuous probability distributions. In our case, we measure the distance between a nominal distribution of the CSPBN and its perturbed version, which corresponds to a true distribution. More specifically, we address the performance of control strategies when the fixed points (or attractors) of constituent BNs of a corresponding CSPBN are perturbed. The attractors are important since they carry the most of the

doi: 10.1049/iet-syb.2008.0121

mass of the steady-state probability distribution of the CSPBN. We show that perturbation of nominal attractors causes a deterioration of the performance of a risk-neutral controller compared to the performance of the risk-sensitive controller, which remains steady even when the majority of attractors are perturbed. Thus, even for a severe discrepancy between the true and the nominal model, the performance of the risk-sensitive controller remains acceptable, which is not the case for a risk-neutral one. Further, we show that by using the same relation between the free-energy and relative entropy, we are able to estimate the influence of a particular attractor on the robustness of the optimal controller. The effectiveness of the proposed algorithm is tested in the case of the GRN related to a malignant melanoma, which has been studied in [21].

The rest of the paper is organised as follows. In Section 2, the CSPBN model is defined. In Section 3, a Markov chain equivalent of the CSPBN is introduced. In Section 4, we formulate the risk-sensitive optimal control problem for the CSPBN and give the solution. In Section 5, an estimation of performance of the risk-sensitive and the risk-neutral optimal control algorithms is provided for the malignant melanoma GRN. Section 6 contains concluding remarks.

1.1 Notation

Here $\boldsymbol{B}$ denotes $\{0, 1\}$, $\boldsymbol{N}$ denotes the set of natural numbers and $\boldsymbol{N}_0$ denotes the set of natural numbers including zero. Here $\boldsymbol{R}$ denotes the set of real numbers, $\boldsymbol{A}^{\mathrm{T}}$ denotes the transpose of matrix $\boldsymbol{A}$ and $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ denotes the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Here $E[X]$ denotes the expected value of a random variable X, and $\oplus$ denotes modulo 2 summation.

2 PBN model

Consider a PBN containing n genes. The activities of genes at a time moment k are given by a binary random vector $\boldsymbol{G}_k = [G_k^1 \cdots G_k^n]^{\mathrm{T}}$, where $G_k^i \in \boldsymbol{B}$ is a scalar binary random variable representing the activity of one gene, where $1 \leq i \leq n$, $k \in \boldsymbol{N}_0$. When G_k^i is 1, a corresponding gene is said to be expressed, while when G_k^i is 0, a corresponding gene is not expressed. The set $\boldsymbol{B}^n$ is called the state space of a PBN.

Remark 1: A point in the state space $\boldsymbol{B}^n$ is called a gene activity pattern (GAP) (or state) of the PBN [3].

A PBN model of the GRN makes a few assumptions with respect to the dynamics of the model. The first is that all genes change their expressions simultaneously, and the other is that the current GAP depends only on the GAP at the previous moment (assumption known as Markovian property). Both assumptions are not true in reality, but are necessary simplifications to make the model computationally feasible. A model that takes into account asynchronous change of expressions of different genes is given in [22].

Having in mind these assumptions, the time evolution of the PBN can be represented by

$$\boldsymbol{G}_{k+1} = \boldsymbol{F}(k, \boldsymbol{G}_k) \tag{1}$$

through a vector-valued mapping $\boldsymbol{F}(k, \boldsymbol{G}_k) = [F^1(k, \boldsymbol{G}_k) \cdots F^n(k, \boldsymbol{G}_k)]^{\mathrm{T}}$. The mappings $F^i(k, \boldsymbol{G}_k)$, $1 \leq i \leq n$, are random Boolean functions, referred to as state transition functions (STFs). In the case of BNs, these functions are deterministic.

Definition 1: If the initial GAP of the GRN is $\boldsymbol{G}_0$, then the sequence of GAPs $\boldsymbol{G}_0, \ldots, \boldsymbol{G}_{\mathrm{T}}$, generated by (1), is called a state trajectory.

The function $\boldsymbol{F}$ maps the previous GAP into the next GAP. When one considers deterministic BNs, the fixed points of the mapping $\boldsymbol{F}$ are called attractors of the BN [4]. From the biological GRN viewpoint, the attractors of a BN correspond to stable states of the GRN. In [3], it is noted that biological GRNs could occupy one of finite number of possible stable states corresponding for instance to growth, differentiation or apoptosis.

For a PBN model, the situation is more complicated. In [4], it is shown that PBN has an equivalent finite state Markov chain representation. The role of attractors in PBNs is played by irreducible subgraphs of the graph describing the Markov chain. When the PBN enters one of these subgraphs, it cannot escape from it, and it is indefinitely captured in that attractor. However, in reality, GRN can escape from a stable state, subject to certain conditions in the cell or its environment. To include this possibility into the model, (1) is adjusted by including an additive noise term

$$\boldsymbol{G}_{k+1} = \boldsymbol{F}(k, \boldsymbol{G}_k) \oplus \boldsymbol{Z}_k \tag{2}$$

where $\boldsymbol{Z}_k \in \boldsymbol{B}^n$ is a disturbance that may be deterministic or random. Here $\boldsymbol{Z}_k$ causes flipping of entries of $\boldsymbol{G}_k$ from 0 to 1 or vice versa. This implies that a corresponding Markov chain is ergodic. If $\boldsymbol{Z}_k$ is a random vector, the flipping of an entry of $\boldsymbol{G}_k$ occurs with some probability p.

Further generalisation of the PBN is called a CSPBN [11]. A CSPBN is a collection of a finite number of deterministic BNs, $S_{\mathrm{BN}} \triangleq \{\mathrm{BN}_1, \mathrm{BN}_2, \ldots, \mathrm{BN}_L\}$. At each time instant, a transition from the current state $\boldsymbol{G}_k$ to the next $\boldsymbol{G}_{k+1}$ is governed by the STFs of a deterministic Boolean network $\mathrm{BN}_i \in S_{\mathrm{BN}}$. At the next moment, the CSPBN may stay in the same network BN_i, or it can change to another $\mathrm{BN}_j \in S_{\mathrm{BN}}$, with some probability q. The choice of a particular BN_j occurs with probability c_j, $1 \leq j \leq L$. If CSPBN ends up in BN_j, the next state $\boldsymbol{G}_{k+2}$ is generated according to BN_j and so on. However, as in the case of the PBN, an equivalent Markov chain can be derived for a CSPBN. A CSPBN reduces to ordinary PBN by setting $q = 1$, implying that in the latter case, the switching of the underlying BN occurs at each time instant. This means that a CSPBN may remain

more time instants in one of the constituent BNs, which better describes the cell processes that are characterised by infrequent switching between stable states [11].

3 Hidden Markov model definition of CSPBNs

In [4], it is shown that a PBN has an equivalent finite-state Markov chain representation having 2^n states. In this section, we formulate a hidden Markov chain model, which will be used in solving a risk-sensitive optimal control problem for partially observed CSPBNs. We treat the case of partially observed CSPBNs, since in reality, as mentioned in the Introduction, many types of noises affect the observation of the GRN state.

Assume that all stochastic signals appearing in the model are defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Denote by $\boldsymbol{X} \triangleq \{\boldsymbol{X}_k\}$, $k \in \boldsymbol{N}_0$, a finite-state Markov chain process representing a PBN and $\boldsymbol{Y} \triangleq \{\boldsymbol{Y}_k\}$, $k \in \boldsymbol{N}$, its noisy observation. Suppose that either $\boldsymbol{X}_0$ or its distribution is known. Further, assume that the state spaces of $\boldsymbol{X}_k \triangleq [X_k^1 \cdots X_k^N]^{\mathrm{T}}$ and $\boldsymbol{Y}_k \triangleq [Y_k^1 \cdots Y_k^M]^{\mathrm{T}}$ are given by

$$S_X \triangleq \{e_1, \ldots, e_N\}, \quad N = 2^n \tag{3}$$

and

$$S_Y \triangleq \{f_1, \ldots, f_M\} \tag{4}$$

respectively. Here, $\boldsymbol{e}_i$, $1 \le i \le N$ (respectively $\boldsymbol{f}_i$, $1 \le i \le M$) are vectors of length N (respectively, M) such that $\boldsymbol{e}_i = [0 \ldots 1 \ldots 0]^{\mathrm{T}}$ (respectively, $\boldsymbol{f}_i = [0 \ldots 1 \ldots 0]^{\mathrm{T}}$), where 1 is the ith entry of the vector. Hence, each of N unit vectors corresponds to one state of the PBN.

Denote by $\{\mathcal{G}_k\}$, $\{\mathcal{F}_k\}$ and $\{\mathcal{Y}_k\}$ complete filtrations generated by $(\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k, \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_k)$, $(X_l, 0 \le l \le k)$ and $(\boldsymbol{Y}_l, 1 \le \boldsymbol{l} \le \boldsymbol{k})$, respectively. A complete filtration of some stochastic process can be understood as its history or the set of all possible past events.

Transitions among Markov chain states represent the transitions among different states of the GRN and are defined by the transition probability matrix $\boldsymbol{A} = (a_{ji}) \in \boldsymbol{R}^{N \times N}$ where

$$a_{ji} \triangleq \Pr\{\boldsymbol{X}_{k+1} = \boldsymbol{e}_j | \boldsymbol{X}_k = \boldsymbol{e}_i\} \tag{5}$$

The Markov property of $\boldsymbol{X}$ dictates

$$\Pr\{\boldsymbol{X}_{k+1} = \boldsymbol{e}_j | \mathcal{F}_k\} = \Pr\{\boldsymbol{X}_{k+1} = \boldsymbol{e}_j | \boldsymbol{X}_k\} \tag{6}$$

Therefore

$$E[\boldsymbol{X}_{k+1} | \mathcal{F}_k] = E[\boldsymbol{X}_{k+1} | \boldsymbol{X}_k] \tag{7}$$

$$= \sum_{i=1}^{N} \boldsymbol{e}_i \Pr\{\boldsymbol{X}_{k+1} = \boldsymbol{e}_i | \boldsymbol{X}_k\} \tag{8}$$

$$= \boldsymbol{A}\boldsymbol{X}_k \tag{9}$$

Further, denote by $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle \triangleq \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$ the inner product of two vectors. Then, since $\boldsymbol{X}_k \triangleq [X_k^1 \ldots X_k^N]^{\mathrm{T}}$

$$X_k^i = \langle \boldsymbol{X}_k, \boldsymbol{e}_i \rangle \tag{10}$$

and

$$\sum_{i=1}^{N} X_k^i = \sum_{i=1}^{N} \langle \boldsymbol{X}_k, \boldsymbol{e}_i \rangle = 1 \tag{11}$$

Next, define

$$a_{k+1}^i = E[\boldsymbol{X}_{k+1}^i | \mathcal{F}_k] \tag{12}$$

$$= E[\langle \boldsymbol{X}_{k+1}, \boldsymbol{e}_i \rangle | \mathcal{F}_k] \tag{13}$$

$$= \sum_{j=1}^{N} \langle \boldsymbol{e}_j, \boldsymbol{e}_i \rangle \Pr\{\boldsymbol{X}_{k+1} = \boldsymbol{e}_j | \boldsymbol{X}_k\} \tag{14}$$

$$= \Pr\{\boldsymbol{X}_{k+1} = \boldsymbol{e}_i | \boldsymbol{X}_k\} \tag{15}$$

and a corresponding vector $\boldsymbol{a}_{k+1} \triangleq [a_{k+1}^1 \cdots a_{k+1}^N]^{\mathrm{T}}$. Here, (14) follows from a Markov property of $\boldsymbol{X}$, and (15) is a consequence of orthogonality of vectors $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ for $i \ne j$.

To describe a time evolution of $\boldsymbol{X}$, a martingale increment $\boldsymbol{V} \triangleq \{\boldsymbol{V}_k\}$, $k \in \boldsymbol{N}$, with respect to $\mathcal{F}_k$ is used. The process V is defined by

$$\boldsymbol{V}_{k+1} \triangleq \boldsymbol{X}_{k+1} - \boldsymbol{A}\boldsymbol{X}_k \tag{16}$$

implying

$$E[\boldsymbol{V}_{k+1} | \mathcal{F}_k] = E[\boldsymbol{X}_{k+1} - \boldsymbol{A}\boldsymbol{X}_k | \boldsymbol{X}_k] \tag{17}$$

$$= \boldsymbol{A}\boldsymbol{X}_k - \boldsymbol{A}\boldsymbol{X}_k \tag{18}$$

$$= 0 \tag{19}$$

Here, (18) is implied by (9). From (19), it follows that $\boldsymbol{V}$ is a martingale increment. Consequently, the time evolution of the Markov chain can be described by

$$\boldsymbol{X}_{k+1} = \boldsymbol{A}\boldsymbol{X}_k + \boldsymbol{V}_{k+1} \tag{20}$$

This also represents the time evolution of the corresponding PBN, and it is known as the state equation.

IET Syst. Biol., 2009, Vol. 3, Iss. 4, pp. 279–295
doi: 10.1049/iet-syb.2008.0121

Further, assume that the noisy observation $\boldsymbol{Y}$ satisfies

$$\Pr\{\boldsymbol{Y}_{k+1} = \boldsymbol{f}_j | \boldsymbol{X}_0, \ldots, \boldsymbol{X}_k, \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_k\} = \Pr\{\boldsymbol{Y}_{k+1} = \boldsymbol{f}_j | \boldsymbol{X}_k\} \quad (21)$$

A transition from a Markov chain state $\boldsymbol{X}_k$ to a noisy observation $\boldsymbol{Y}_{k+1}$ is defined by a transition probability matrix $\boldsymbol{C} = (c_{ji}) \in \boldsymbol{R}^{M \times N}$ where

$$c_{ji} \triangleq \Pr\{\boldsymbol{Y}_{k+1} = \boldsymbol{f}_j | \boldsymbol{X}_k = \boldsymbol{e}_i\} \quad (22)$$

By the same line of reasoning as for the state process $\boldsymbol{X}$, the noisy observation $\boldsymbol{Y}$ satisfies

$$E[\boldsymbol{Y}_{k+1} | \mathcal{G}_k] = \boldsymbol{C}\boldsymbol{X}_k \quad (23)$$

and its time evolution is given by the observation equation

$$\boldsymbol{Y}_{k+1} = \boldsymbol{C}\boldsymbol{X}_k + \boldsymbol{W}_{k+1} \quad (24)$$

$\boldsymbol{W} \triangleq \{\boldsymbol{W}_k\}$, $k \in \boldsymbol{N}$, is a martingale increment with respect to $\mathcal{G}_k$ such that

$$E[\boldsymbol{W}_{k+1} | \mathcal{G}_k] = 0 \quad (25)$$

and $\boldsymbol{W}$ and $\boldsymbol{V}$ are mutually independent. In addition, introduce

$$c^i_{k+1} = E[Y^i_{k+1} | \mathcal{G}_k] \quad (26)$$

$$= \Pr\{\boldsymbol{Y}_{k+1} = \boldsymbol{f}_i | \boldsymbol{X}_k\} \quad (27)$$

and a corresponding vector $\boldsymbol{c}_{k+1} \triangleq [c^1_{k+1} \cdots c^M_{k+1}]^{\mathrm{T}}$.

The CSPBN is completely described by (20) and (24), and next an optimal risk-sensitive control problem shall be defined for such a CSPBN. For further elaboration on (20) and (24) see Elliott *et al.* [18].

4 Risk-sensitive control of CSPBNs

In this section, we consider a risk-sensitive control of CSPBNs for a finite horizon case. Finite horizon control implies that the control signal (or intervention) is applied a finite number of times.

We first define the dynamics of the controlled CSPBN, and then introduce a risk-sensitive cost function and optimal control problem. The optimal control problem is solved by the dynamic programming, where we follow the approach found in [18]. At the end, we discuss the robustness of the risk-sensitive controller with respect to uncertainties coming from the lack of knowledge of the true CSPBN distribution. This represents one of the main contributions of the paper, where the robustness is evaluated by applying the relation between the free-energy and the relative entropy.

4.1 System dynamics

In the case of controlled CSPBNs, the transition probability matrix $\boldsymbol{A}$ becomes a function of control signals that change the activity of targeted genes. Then, the CSPBN is given by

$$\boldsymbol{X}^u_{k+1} = \boldsymbol{A}(\boldsymbol{u}_k)\boldsymbol{X}^u_k + \boldsymbol{V}_{k+1} \quad (28)$$

$$\boldsymbol{Y}_{k+1} = \boldsymbol{C}\boldsymbol{X}^u_k + \boldsymbol{W}_{k+1} \quad (29)$$

As noticed in the Introduction, the overall dynamics of the CSPBN is now defined not only by internal dynamics (20), but also by external control. All signals are now defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P}^u)$. A control sequence is denoted by $\boldsymbol{u} \triangleq \{\boldsymbol{u}_k\}$, where $\boldsymbol{u}$ takes values in some measurable space U, and $\boldsymbol{u}_k$ is $\mathcal{Y}_k$-measurable. Denote a set of such controls by $U(k)$. Then, $U(k, k+l)$ is used to denote $U(k) \cup U(k+1) \cup \cdots \cup U(k+l)$, where $[\boldsymbol{u}_0 \cdots \boldsymbol{u}_{K-1}]^{\mathrm{T}} \in U(0, K-1)$, and K defines a finite horizon.

The control problem goal is to find a control sequence $\boldsymbol{u}^o$ that minimises a risk-sensitive cost that is introduced next. The optimal control sequence $\boldsymbol{u}^o$ is applied by the controller, which makes its decision based on the observation $\boldsymbol{Y}$.

4.2 Risk-sensitive cost function and definition of control problem

A risk-sensitive cost function, which depends on the control signal $\boldsymbol{u}$ and the state $\boldsymbol{X}^u$ is given by

$$J^{\mu}(\boldsymbol{X}_0, \boldsymbol{u}) \triangleq \exp\left(\mu\left\{\sum_{k=0}^{K-1} L(\boldsymbol{X}^u_k, \boldsymbol{u}_k) + \Phi(\boldsymbol{X}^u_K)\right\}\right) \quad (30)$$

where $\mu > 0$ is a risk-sensitive parameter, and $L(\boldsymbol{X}^u_k, \boldsymbol{u}_k)$ is the cost at the moment k, which depends on the current state $\boldsymbol{X}^u_k$ and the control signal $\boldsymbol{u}_k$. The function Φ defines the terminal cost. When the distribution of $\boldsymbol{X}_0$ is given by π_0, the average risk-sensitive cost becomes

$$J^{\mu}_1(\pi_0, \boldsymbol{u}) \triangleq E[J^{\mu}(\boldsymbol{X}_0, \boldsymbol{u})] \quad (31)$$

The risk-sensitive stochastic control problem for the CSPBN is defined by

$$J^{\mu}_1(\pi_0, \boldsymbol{u}^o) = \inf_{\boldsymbol{u} \in U(0,K-1)} J^{\mu}_1(\pi_0, \boldsymbol{u}) \quad (32)$$

The cost function J^{μ}_1 is usually defined such that the undesirable states of the GRN admit larger cost then desirable ones; similarly, the application of non-zero control incurs larger cost as well. Hence, the optimal sequence $\boldsymbol{u}^o$ minimises the possibility of undesirable state occurrence and excessive use of the control.

4.3 Information-state

When the state process $\boldsymbol{X}^u$ is completely observable, a controller uses this information to minimise a cost function. When the state process is only partially observable, like in our case, a

controller has to be provided with the state estimation. In the case of risk-sensitive control, a notion of information-state is used instead [18]. In addition to the state process, the information-state includes part of the exponential cost function. The information-state gives the complete information about the state process $\boldsymbol{X}^u$ available in the measurement data. Our goal is to find a recursive formula for the evolution of the information-state in the time domain as in the case of the Kalman filter.

To find the recursive formula and solve optimal control problem, we employ the change-of-measure technique proposed in [18]. The meaning of the change-of-measure is that a system originally defined on a probability measure $\mathcal{P}$ can be transformed into another system that is defined on a probability measure $\bar{\mathcal{P}}$. However, under the new measure $\bar{\mathcal{P}}$, some or all stochastic processes become i.i.d. Then, all computations are performed under this ideal measure. Afterwards, by an inverse change-of-measure, the results can be interpreted under the original measure $\mathcal{P}$.

Next, we introduce the information-state, denoted by $\boldsymbol{q}_k^\mu \triangleq [q_k^\mu(1) \cdots q_k^\mu(N)]^{\mathrm{T}} \in \boldsymbol{R}^N$, by

$$\langle \boldsymbol{q}_k^\mu, \boldsymbol{e}_i \rangle = q_k^\mu(i) \triangleq \bar{E}\left[I(\boldsymbol{X}_k^u = \boldsymbol{e}_i)\bar{\Lambda}_k \exp\left(\mu \sum_{l=0}^{k-1} L(\boldsymbol{X}_l^u, \boldsymbol{u}_l) \right) \middle| \mathcal{Y}_k \right], \quad 1 \le i \le N \tag{33}$$

where $I(\boldsymbol{X}_k^u = \boldsymbol{e}_i)$ is the indicator function, equal to one when $\boldsymbol{X}_k^u = \boldsymbol{e}_i$ and zero otherwise, and $\bar{\Lambda}_k$ denotes a Radon–Nykodim derivative, defining the inverse change-of-measure, that is given in the Appendix by (65). Here, $\bar{E}$ denotes the expectation with respect to a new probability measure $\bar{\mathcal{P}}$ under which $\{\boldsymbol{Y}_k\}$ are i.i.d. and have uniform distribution over S_Y. For a detailed explanation of the change-of-measure technique see Appendix.

The recursive formula for the time evolution of the information-state is given in the following theorem.

Theorem 1: The time evolution of the information-state when $M = N$ is given by

$$\boldsymbol{q}_{k+1}^\mu = \Sigma^{\mu *}(\boldsymbol{u}_k, \boldsymbol{Y}_{k+1})\boldsymbol{q}_k^\mu \tag{34}$$

where the operator $\Sigma^{\mu *}(\boldsymbol{u}_k, \boldsymbol{Y}_{k+1})$, having operand $\boldsymbol{q}_k^\mu$, is defined by

$$\Sigma^{\mu *}(\boldsymbol{u}_k, \boldsymbol{Y}_{k+1})q^\mu(j) \triangleq \sum_{i=1}^{N} \Pr\{\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j | \boldsymbol{X}_k^u = \boldsymbol{e}_i\} \times \exp\big(\mu L(\boldsymbol{e}_i, \boldsymbol{u}_k)\big) \Pr\{\boldsymbol{Y}_{k+1} | \boldsymbol{X}_k^u = \boldsymbol{e}_i\} q_k^\mu(i), \quad 1 \le j \le N \tag{35}$$

or in more compact matrix form

$$\Sigma^{\mu *} = \boldsymbol{A}\Upsilon \tag{36}$$

where Υ is a diagonal matrix given by $\mathrm{diag}[\Pr\{\boldsymbol{Y}_{k+1} | \boldsymbol{X}_k = e_1\} \mathrm{e}^{\mu L(\boldsymbol{e}_1, \boldsymbol{u}_k)} \cdots \Pr\{\boldsymbol{Y}_{k+1} | \boldsymbol{X}_k = \boldsymbol{e}_N\} \mathrm{e}^{\mu L(\boldsymbol{e}_N, \boldsymbol{u}_k)}]$.

Proof: The proof is given in the Appendix.

The cost function (31) can also be represented by using the change-of-measure in the following way

$$J_1^\mu(\pi_0, \boldsymbol{u}) = \bar{E}\left[\bar{\Lambda}_K \exp\left(\mu \left\{ \sum_{k=0}^{K-1} L(\boldsymbol{X}_k^u, \boldsymbol{u}_k) + \Phi(\boldsymbol{X}_K^u) \right\} \right) \right] \tag{37}$$

Further, by employing the definition of the information-state, the cost function can be expressed by

$$J_1^\mu(\pi_0, \boldsymbol{u}) = \bar{E}[\langle \boldsymbol{q}_K^\mu, \mathbf{exp}(\boldsymbol{\mu\Phi}) \rangle] \tag{38}$$

where $\mathbf{exp}(\boldsymbol{\mu\Phi}) \triangleq [\mathrm{e}^{\mu\Phi(\boldsymbol{e}_1)} \cdots \mathrm{e}^{\mu\Phi(\boldsymbol{e}_N)}]^{\mathrm{T}}$. This follows from

$$\langle \boldsymbol{q}_K^\mu, \mathbf{exp}(\boldsymbol{\mu\Phi}) \rangle = \sum_{i=1}^{N} q_K^\mu(i) \exp(\mu\Phi(\boldsymbol{e}_i)) \tag{39}$$

$$= \sum_{i=1}^{N} \exp(\mu\Phi(\boldsymbol{e}_i)) \bar{E}\left[I(\boldsymbol{X}_K^u = \boldsymbol{e}_i)\bar{\Lambda}_K \times \exp\left(\mu \sum_{k=0}^{K-1} L(\boldsymbol{X}_k^u, \boldsymbol{u}_k) \right) | \mathcal{Y}_K \right] \tag{40}$$

$$= \bar{E}\left[\sum_{i=1}^{N} \exp(\mu\Phi(\boldsymbol{e}_i)) I(\boldsymbol{X}_K^u = \boldsymbol{e}_i)\bar{\Lambda}_K \times \exp\left(\mu \sum_{k=0}^{K-1} L(\boldsymbol{X}_k^u, \boldsymbol{u}_k) \right) | \mathcal{Y}_K \right] \tag{41}$$

$$= \bar{E}\left[\exp(\mu\Phi(\boldsymbol{X}_K^u))\bar{\Lambda}_K \times \exp\left(\mu \sum_{k=0}^{K-1} L(\boldsymbol{X}_k^u, \boldsymbol{u}_k) \right) | \mathcal{Y}_K \right] \tag{42}$$

Here, (42) follows because $\boldsymbol{X}_K$ is equal to one of N vectors $\boldsymbol{e}_i$ for which the indicator function $I(\boldsymbol{X}_K^u = \boldsymbol{e}_i)$ is equal to 1. Taking the expectation $\bar{E}$ of (42) with respect to $\mathcal{Y}_K$, the cost function with respect to the probability measure $\bar{\mathcal{P}}$, (37), is obtained, concluding the proof of (38).

4.4 Optimal control by dynamic programming

The solution of the risk-sensitive optimal control problem applies the so-called backward-forward principle, since the risk-sensitive cost (38) can be expressed in terms of forward and backward processes. The forward process is given by

doi: 10.1049/iet-syb.2008.0121

the information-state recursion (34). For backward process, we need to introduced the adjoint operator Σ^{μ} of $\Sigma^{\mu*}$ such that

$$\langle \Sigma^{\mu*}\tau,\ \eta\rangle = \langle \tau,\ \Sigma^{\mu}\eta\rangle \tag{43}$$

where

$$\Sigma^{\mu}(\boldsymbol{u}_k,\ \boldsymbol{Y}_{k+1})\beta(i) = N\exp\big(\mu L(\boldsymbol{e}_i,\ \boldsymbol{u}_k)\big)\Pr\{\boldsymbol{Y}_{k+1}|\boldsymbol{X}_k = \boldsymbol{e}_i\}\sum_{j=1}^{N}\Pr\{\boldsymbol{X}^{u}_{k+1} = \boldsymbol{e}_j|\boldsymbol{X}^{u}_{k} = \boldsymbol{e}_i\}\beta(j),\quad 1\le i\le N$$

or expressed in a compact matrix form

$$\Sigma^{\mu} \triangleq \Upsilon A^{\mathrm{T}} \tag{44}$$

The backward process is now defined as

$$\boldsymbol{\beta}^{\mu}_{k} = \Sigma^{\mu}(\boldsymbol{u}_k,\ \boldsymbol{Y}_{k+1})\boldsymbol{\beta}^{\mu}_{k+1} \tag{45}$$

$$\boldsymbol{\beta}^{\mu}_{K} = \mathbf{exp}(\boldsymbol{\mu\Phi}) \tag{46}$$

The dynamic programming for the solution of the risk-sensitive control problem relies on the following lemma that relates the inner product of the forward and backward processes at different time instants.

Lemma 1: $\langle \boldsymbol{q}^{\mu}_{k},\ \boldsymbol{\beta}^{\mu}_{k}\rangle = \langle \boldsymbol{q}^{\mu}_{k-1},\ \boldsymbol{\beta}^{\mu}_{k-1}\rangle$.

Proof:

$$\langle \boldsymbol{q}^{\mu}_{k},\ \boldsymbol{\beta}^{\mu}_{k}\rangle = \langle \Sigma^{\mu*}\boldsymbol{q}^{\mu}_{k-1},\ \boldsymbol{\beta}^{\mu}_{k}\rangle \tag{47}$$

$$= \langle \boldsymbol{q}^{\mu}_{k-1},\ \Sigma^{\mu}\boldsymbol{\beta}^{\mu}_{k}\rangle \tag{48}$$

$$= \langle \boldsymbol{q}^{\mu}_{k-1},\ \boldsymbol{\beta}^{\mu}_{k-1}\rangle \tag{49}$$

where (47) follows from Theorem 1, (48) from the adjoint operator property (43), and (49) from (45) □

Next, by employing the forward and backward processes, the control problem equivalent to the one given in (32) is introduced. Assume that the information-state is known at the moment k, $\boldsymbol{q}^{\mu}_{k} = \boldsymbol{q}$. By the recursive relation (34), starting from the moment k, the information-state is available till the finite horizon K. Then, the cost function from the moment k till the finite horizon K is defined by

$$J^{\mu}_{1}(k,\ \boldsymbol{q}) = \inf_{\boldsymbol{u}\in U(k,K-1)}\bar{E}[\langle \boldsymbol{q}^{\mu}_{k},\ \boldsymbol{\beta}^{\mu}_{k}\rangle|\boldsymbol{q}^{\mu}_{k} = \boldsymbol{q}] \tag{50}$$

$$= \inf_{\boldsymbol{u}\in U(k,K-1)}\bar{E}[\langle \boldsymbol{q}^{\mu}_{K},\ \mathbf{exp}(\boldsymbol{\mu\Phi})\rangle] \tag{51}$$

The equality in (51) is implied by Lemma 1 by noticing that $\langle \boldsymbol{q}^{\mu}_{k},\ \boldsymbol{\beta}^{\mu}_{k}\rangle = \langle \boldsymbol{q}^{\mu}_{k+1},\ \boldsymbol{\beta}^{\mu}_{k+1}\rangle = \cdots\langle \boldsymbol{q}^{\mu}_{K},\ \boldsymbol{\beta}^{\mu}_{K}\rangle$, and (46). The solution of the risk-sensitive control problem (51) is given by the following theorem.

Theorem 2: (Dynamic programming for the risk-sensitive control of CSPBNs) The cost function of the risk-sensitive control problem defined by (51) satisfies the following recursion

$$J^{\mu}_{1}(k,\ \boldsymbol{q}) = \inf_{\boldsymbol{u}\in U(k,k)}\bar{E}\big[J^{\mu}_{1}(k+1,\ \Sigma^{\mu*}(\boldsymbol{u}_k,\ \boldsymbol{Y}_{k+1})\boldsymbol{q})\big]$$

$$J^{\mu}_{1}(K,\ \boldsymbol{q}) = \langle \boldsymbol{q}^{\mu}_{K},\ \mathbf{exp}(\boldsymbol{\mu\Phi})\rangle \tag{52}$$

Proof: The proof is along the same lines as in the [18].

Remark 2: The significance of the previous theorem is that it gives a recursive solution for the control problem defined by (32), making it computationally tractable. In Section 5, this recursive solution is applied for the control of malignant melanoma GRN, where each step of the control is explained in more detail.

4.5 Robustness of risk-sensitive controller

There are different ways to characterise the robustness of the risk-sensitive controllers. One is to apply the small-noise limit that shows the equivalence with H^{∞} robust control [18]. The other method, which we use, employs the duality between the free-energy and the relative entropy (see [23]). Assume first that there exists a measurable function Ψ bounded from below. Then the free-energy can be expressed by means of the relative entropy $D(Q\|P)$ as

$$\log\int\exp(\mu\Psi)\mathrm{d}P = \sup_{Q}\left\{\int\mu\Psi dQ - D(Q||P):\ D(Q||P)<\infty\right\}$$

Hence, the free-energy represents the log of the mean exponential cost, and it corresponds to the log risk-sensitive cost defined by (31). Let us adjust the duality relation to our problem as follows

$$\log\int\exp(\mu\Psi(X^{u},\ u))\mathrm{d}\mathcal{P}^{u} = \sup_{Q\in\mathcal{Q}}\int\mu\Psi(X^{u},\ u)\mathrm{d}Q - D(Q||\mathcal{P}^{u}) \tag{53}$$

where $\mathcal{Q} \triangleq \{Q : D(Q||\mathcal{P}^{u})<\infty\}$, and Ψ represents the exponent in (30). The distribution $\mathcal{P}^{u}$ plays the role of the nominal, while Q is the true although unknown distribution of the CSPBN. By appropriately transforming (53), we obtain an upper bound on the true (under the true CSPBN distribution Q) non-exponential (i.e. risk-neutral) cost Ψ

$$\int\Psi(X^{u},\ u)\mathrm{d}Q \le \frac{1}{\mu}\log\int\exp(\mu\Psi(X^{u},\ u))\mathrm{d}\mathcal{P}^{u} + \frac{1}{\mu}D(Q||\mathcal{P}^{u}) \tag{54}$$

The upper bound is given in terms of the free-energy, that is, the log of the risk-sensitive cost under the nominal distribution of the CSPBN $\mathcal{P}^{u}$, and the relative entropy $D(Q\|\mathcal{P}^{u})$ between

the true distribution and the nominal one. Therefore (54) can be used to assess the performance of the risk-sensitive controller under an assumed true CSPBN distribution Q and the nominal distribution $\mathcal{P}^u$ obtained through the measurements.

The performance of the controller depends on the distance between the true and the nominal distribution, expressed by the relative entropy $D(Q\|\mathcal{P}^u)$, and the risk-sensitive parameter μ. The more accurate CSPBN model the better the performance of the controller. From the point of view of attractors being one of the most important parameters that define the dynamics of the CSPBNs, the more accurate knowledge of attractors of the CSPBN, the better the performance of the controller.

A value of the risk-sensitive parameter μ depends on application. It essentially affects the performance of the controller, that is, its ability to cope with the uncertainty. The choice of the risk-sensitive parameter μ will be discussed in Section 5.

Next, (53) is further transformed to get even deeper insight into the merit of the risk-sensitive control. Since the relative entropy is finite, we may assume it is bounded by some finite constant R, $D(Q\|\mathcal{P}^u) \leq R$. A modified constraint set $\mathcal{Q}_m$ is given by

$$\mathcal{Q}_m \triangleq \{Q : D(Q\|\mathcal{P}^u) \leq R\} \tag{55}$$

Then, we have the following theorem.

Theorem 3: For a given u, which belongs to a set of admissible controls U_{ad} and some $s \in \boldsymbol{R}$ such that Ψ/s is a measurable function bounded from below, the following holds

1.

$$\sup_{Q\in\mathcal{Q}_m} \int \Psi(X^u, u)\mathrm{d}Q = \inf_{s\geq 0}\left\{ s \log \int \exp\left(\frac{\Psi(X^u, u)}{s}\right)\mathrm{d}\mathcal{P}^u + sR \right\} \tag{56}$$

2.

$$\lim_{s\to\infty} s \log \int \exp\left(\frac{\Psi(X^u, u)}{s}\right)\mathrm{d}\mathcal{P}^u = \int \Psi(X^u, u)\mathrm{d}\mathcal{P}^u \tag{57}$$

Proof: The proof may be found in [23].

There are two major implications of the previous theorem:

- Statement (1) of Theorem 3 explains why a minimisation of the risk-sensitive cost [the right-hand side of (56)] over the set of admissible controls U_{ad} provides a robust controller; a minimisation of the risk-sensitive cost corresponds to the minimisation of the worst-case cost, that is, the maximal risk-neutral cost [the left-hand side of (56)], where the maximisation is performed over the set of all possible CSPBNs' distributions that are at the distance at most R from the nominal distribution $\mathcal{P}^u$. In the robust control theory, the set $\mathcal{Q}_m$ is called the uncertainty set. The parameter R defines the size of the uncertainty set, and it is determined by the quality of a CSPBN model. The more accurate CSPBN model, the smaller R;

- Statement (2) of Theorem 3 shows that the risk-neutral cost is recovered from the risk-sensitive cost by letting s go to infinity. By comparing (53) with (57), it can be noticed that s corresponds to the reciprocal risk-sensitive parameter μ. Thus, when μ tends to zero, the risk-sensitive controller becomes risk-neutral under the nominal distribution $\mathcal{P}^u$. This statement is used in the next section to compare risk-sensitive controllers having different values of the parameter μ.

5 Risk-sensitive controller performance estimation

The robustness of risk-sensitive control shall be illustrated for an example of a CSPBN designed from expression data in the case of a malignant melanoma found in [21, 24, 25].

5.1 Background in expression profile study of malignant melanoma

In [21], gene expression profiles were used to distinguish between different malignant melanoma samples that are otherwise indistinguishable by histopathological, molecular or immunohistochemical markers. This study gathered expression profiles from 38 samples, of which 31 were cutaneous melanomas and 7 were control samples. This messenger RNA profiling analysis revealed that there were 6791 unique genes in melanoma samples. By applying different clustering techniques to the 31 melanoma samples, two clusters were obtained, one containing 19 and other containing 12 samples. The first cluster was highly compact showing overall similarity of expression profiles over 19 samples. The validity of clustering predictions based on the 19 element major cluster was evaluated by further statistical tests confirming that this major cluster is a good representative of studied melanoma samples. For each gene, a discriminative weight was computed that assesses the ability of the expression of a gene to discriminate between two clusters. On the basis of this clustering analysis, it was found that many genes underlying this 19 element cluster were also differentially regulated in invasive melanomas. Thus, expression of these genes may serve as the early markers signalling the differentiation of cells into invasive melanomas. Among the expressed genes, WNTA5 has the largest discriminative weight implying that this gene has a potential to be a strong marker of aggressive behaviour of cutaneous melanoma.

The following study [24] further explored the role of Wnt5a protein in malignant melanoma. Constitutive expression of Wnt5a led to an increase in motility and invasion in the

doi: 10.1049/iet-syb.2008.0121

melanoma cells. In [24], the authors also examined the molecular changes that happened in invasive melanoma cells upon Wnt5a expression. It was shown that only in Wnt5a-transfected cells the level of activated protein kinase C (PKC) increased significantly, and there was a direct correlation between Wnt5a expression and increased invasion and motility of the melanoma cells. Consistent with these findings, treating the cells with antibodies to Frizzled-5, a Wnt5a receptor, resulted in reduced motility and invasion of melanoma cells perhaps by disrupting the Wnt5a signalling to PKC.

The above consideration supports an idea that is possible to design therapies (based on the external control of GRNs) for a treatment of the malignant melanoma by suppressing the activity of Wnt5a gene. More precisely, whenever the level of Wnt5a in cells is high, that potentially corresponds to malignant melanoma, a therapy could consist of inhibiting Wnt5a activity. This should in turn disrupt the signal pathway between Wtn5a and PKC (Fig. 1) resulting in reduced melanoma invasiveness.

5.2 Robust control of malignant melanoma

The optimal external control of a GRN requires a suitable mathematical model. That is why in [8, 25], the gene expression data found in [21, 24] are used to extract a group of seven genes that best represent melanoma samples. Here, 'best' means that these seven genes have an ability to efficiently cross-predict each other. The extracted genes are WNT5A, pirin, S100P, RET1, MART1, HADHB and STC2.

In [8], a PBN model of the GRN has been designed consisting of previous seven genes, and the dynamic programming is used to find optimal strategies for controlling malignant melanoma. Further evolution of a PBN model is a CSPBN model introduced in [11]. For the same seven genes, a CSPBN is built, and a finite and infinite horizon control problems are studied in [11, 12], respectively. The control strategies given in these two papers optimise the so-called risk-neutral cost functions.

In this section, we present the results on the robust risk-sensitive control of the CSPBN found in [12], when the model uncertainty is described through the uncertainty of the probability distribution of the CSPBN. In the context of

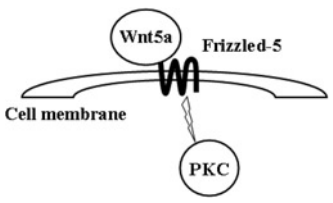

Figure 1 *Signal-transduction pathway between Wnt5a and PKC*

CSPBNs defined by (2), the GAP $\boldsymbol{G}_k = [G_k^1, \ldots, G_k^7]^{\mathrm{T}}$ of the seven gene CSPBN is defined by the activity of WNT5A (represented by G_k^1), pirin (G_k^2), S100P (G_k^3), RET1 (G_k^4), MART1 (G_k^5), HADHB (G_k^6) and STC2 (G_k^7). If the GAP $\boldsymbol{G}_k$ is understood as a binary word, WNT5A is not expressed for the first 64 GAPs, while for the next 64 GAPs WNT5A is expressed. It should be noticed that this CSPBN has been designed by the algorithm given in [26], which combines four most probable deterministic BNs each having different subsets of singleton attractors. The transition probabilities of the CSPBN are obtained by averaging the transition probabilities of constituent BNs.

Example 1: Next, we give an example of the risk-sensitive control of the CSPBN for a sample path of the noisy observation Y. The number of states of the CSPBN is $N = 2^7 = 128$. The states are unit-vectors coming from the set $S_X = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{128}\}$. The value of the finite horizon K, which determines the number of time instants the control is applied is here arbitrarily set to 5; however, in practice it should be determined with the help of physician.

It is further assumed that: (1) $S_X \equiv S_Y$, (2) Transition probabilities c_{ji}, $1 \leq i, j \leq N$ defined by (22), do not depend on the control u, $c_{ji} = p_t$ for $j = i$ and $c_{ji} = 1 - p_t/(N-1)$ for $j \neq i$, (3) The control signal u is binary such that $u_k = 0$ means that the control is not exercised, while $u_k = 1$ means the control is exercised.

The second assumption implies that the overall noise impeding the microarray device is independent on the applied control u.

From the third assumption, if $u_k = 0$, the CSPBN evolves according to (20). Also, the control $u_k = 1$ can only be applied when the gene WNT5A is expressed in which case the activity of WNT5A is forcefully inhibited by the external input. The transition probability matrices $\boldsymbol{A}(u_k)$ for $u_k = 0$ and $u_k = 1$ are generated by using the approach found in [11] and in accordance with the above discussion.

The function L, which defines the cost function is chosen as follows

$$L(\boldsymbol{X}_k^u, u_k) \triangleq \langle \boldsymbol{\mathcal{R}}_k, \boldsymbol{X}_k^u \rangle + \langle \boldsymbol{\mathcal{B}}_k, u_k \rangle \tag{58}$$

Since $\boldsymbol{X}_k^u = \boldsymbol{e}_j$, $1 \leq j \leq 128$, the ith element of the vector $\boldsymbol{\mathcal{R}}_k$ determines the cost of the ith CSPBN state at moment k. Similarly, the cost of the control at moment k, u_k, is determined by the vector $\boldsymbol{\mathcal{B}}_k$. Whether the control is carried out or not depends on the optimal control algorithm. Following the choice of the parameters in [8]

$$L(\boldsymbol{X}_k^u, u_k) = \begin{cases} \langle \boldsymbol{\mathcal{R}}_k, \boldsymbol{X}_k^u \rangle + \langle \boldsymbol{\mathcal{B}}_k, u_k \rangle & k \neq 5 \\ \boldsymbol{\Phi}(\boldsymbol{X}_k^u) = \langle \boldsymbol{\mathcal{R}}_k, \boldsymbol{X}_k^u \rangle & k = 5 \end{cases}$$

where the first 64 entries of $\mathcal{R}_k$ and $\mathcal{B}_k$ are zeros and the last 64 are set to 10. Thus, the cost function penalises the controls and

states for all time instants, while only the state X_k^u is penalised for the last moment $k = 5$ when WNT5A is expressed. In practice, the entries of $\mathcal{R}_k$ and $\mathcal{B}_k$ should be determined by a physician.

The value of p_t is set to 0.99, and $\mu = 0.7$.

The initial information-state is given by

$$q_0^{0.7} = \Pr\{X_0\} = \begin{cases} 1 & X_0 = 41 \\ 0 & X_0 \neq 41 \end{cases}$$

implying that the initial state is set to 41.

A noisy observation is given by $Y_1 = 15$, $Y_2 = 1$, $Y_3 = 87$, $Y_4 = 5$, $Y_5 = 1$. This sample path is selected since the transition probability matrix A of the uncontrolled CSPBN suggests that this sample path is highly probable. Because of this and because the noise level is low ($p_t = 0.99$), it can be predicted that the control will be applied only at $k = 2$.

For such defined parameters, the optimal control signal is computed as follows. Given the observation Y, the information-state $\{q_k^\mu\}$ is calculated by using the recursion (34). This procedure is illustrated in Fig. 2 by a tree. The root of the tree represents the initial information-state $q_0^{0.7}$. The next level corresponds to the possible information-states at the moment $k = 1$ depending on the value of the control signal u_0. In general, a current state at the moment k is branching towards two possible information-states depending on the control signal u_k. The depth of the tree is determined by the finite horizon $K = 5$. It can be seen that the size of the tree grows exponentially with the finite horizon K. The optimal control is found by applying Theorem 2, that is, backtracking of the tree diagram looking for the minimal cost. Bold branches in the tree define the optimal path, and the optimal control sequence is the one obtained by reading the labels along the branches starting from the root node. As correctly predicted, the control $u_k = 1$ is applied at the moment $k = 2$; otherwise $u_k = 0$. The cost, when the optimal control is applied, is equal to 4.48176×10^{-6}; the cost without control is 1.29989×10^{-4}.

Example 2: The second example determines the optimal control in the case of a randomly chosen sample path $Y_1 = 41$, $Y_2 = 35$, $Y_3 = 62$, $Y_4 = 4$, $Y_5 = 97$. We are allowed to pick a random path because of the ergodicity of the CSPBN. The value of p_t is set to 0.9, and $\mu = 0.5$. The initial information-state $q_0^{0.5}$ is chosen similarly to the one given in Example 1 so that $X_0 = 10$. The cost function L remains the same as in Example 1.

A tree representing the computation of the optimal control is shown in Fig. 3. Since the sample path is not highly probable, now it is not possible to predict correctly the optimal control sequence u. At this point, it should be realised that the information-state plays the role of a filter that provides the controller with reliable information about the sequence of states of the CSPBN such that the controller can make a right decision regardless of the noisy state observations. The optimal control results in the cost of 3.17336×10^{-5}, while the cost in the absence of control is an order of magnitude worse and is equal to 3.48849×10^{-4} .

5.3 Risk-sensitive against risk-neutral control

To select a right controller, we need to make a fair comparison of their performances. From the point of view of robust performance comparison, the upper bound given by (54) is useful since it relates the risk-sensitive cost under the nominal probability distribution of the CSPBN, which is minimised, and corresponding risk-neutral cost under the true probability distribution. The risk-neutral cost under the true probability distribution is a true cost experienced by the system. Therefore by using (54), we are able to compare the robustness of controllers for different risk-sensitive parameters

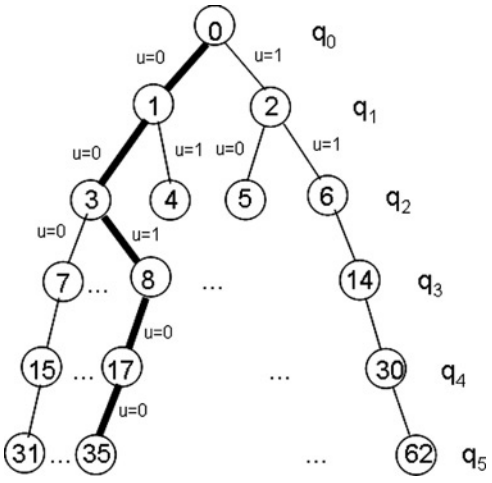

Figure 2 *Information-state tree for the risk-sensitive control for the first noisy observation sample path*

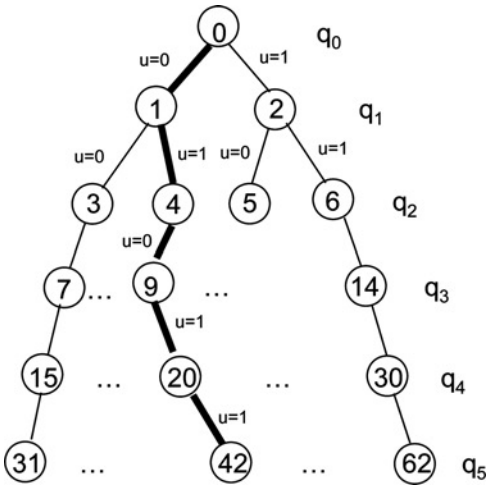

Figure 3 *Information-state tree for the risk-sensitive control for the second noisy observation sample path*

doi: 10.1049/iet-syb.2008.0121

μ, by assuming different true distributions and observing how the upper bound on the true cost varies.

Candidates for true probability distributions shall be constructed from the nominal distribution. This is based on the assumption that the nominal distribution obtained through the measurements is in vicinity of the true distribution. We construct a true distribution by changing the probabilities of the nominal attractors. The reason for focusing on attractors is that the choice of attractors and their probabilities determines the validity of the CSPBN model. Here, when we say 'attractor', we mean an attractor of constituent deterministic BN of corresponding CSPBN. The attractors of the CSPBN found in [12] are the states: 22, 28, 46, 52, 54, 62, 63, 65, 72, 73, 79, 81 and 104.

We change the jth attractor probability by changing its transition probabilities a_{ij}, $1 \leq i \leq 128$. The probability of transition into itself a_{jj} is reduced by some percentage (in our examples this portion is set to 99.9%) and that amount is equally distributed among the rest of transition probabilities a_{ij}, $i \neq j$. In this way, the probability of staying in the attractor is reduced, while the probabilities of transitions into neighbouring states is increased. Consequently, the network spends less time in the attractor reducing its importance.

We start from the attractor 22, and compute the upper bound (54) when the probability of only that attractor has been changed. In the next step, we keep the probability of the attractor 22 changed, and in addition change the probability of the attractor 28. The upper bound for this case is computed as well. We proceed further till we change the probabilities of all attractors. The graphs of upper bounds against the number of changed attractors are shown in Figs. 4–6. Each integer point on the abscissa represents one possible true distribution of the CSPBN. Fig. 4 corresponds to the initial state 0, Fig. 5 to 10, and Fig. 6 to 70. In addition, Fig. 7 shows the variation of the relative entropy between the nominal and the true distribution of the CSPBN against number of neutralised attractors for different initial states.

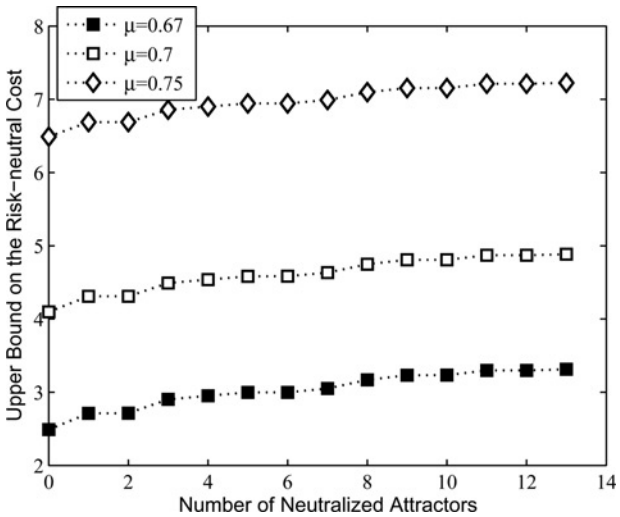

Figure 4 *Upper bound on the risk-neutral cost under the true distribution against the number of neutralised attractors where the initial state is set to 0*

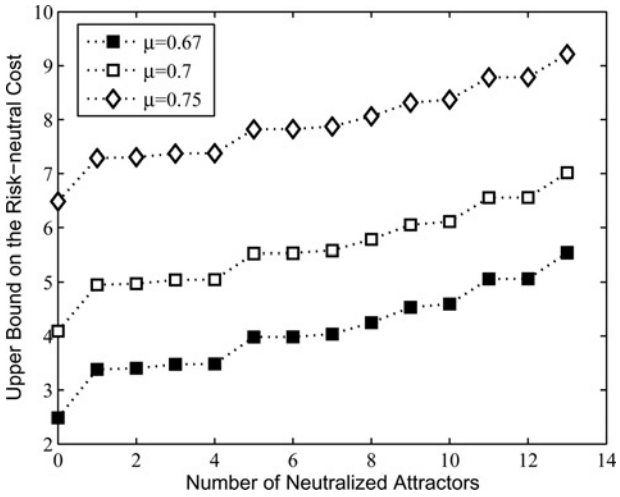

Figure 5 *Upper bound on the risk-neutral cost under the true distribution against the number of neutralised attractors where the initial state is set to 10*

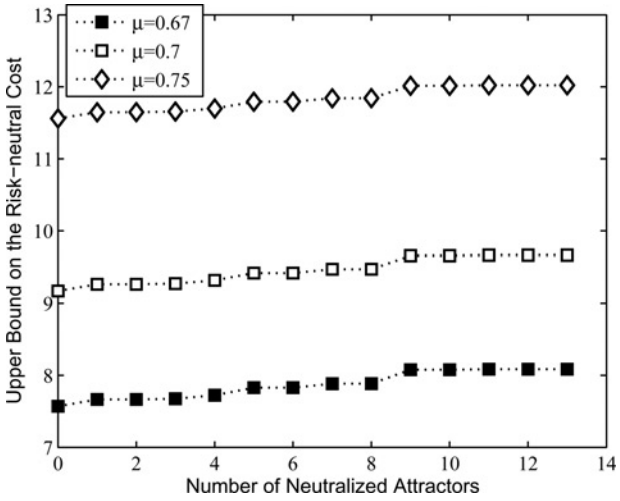

Figure 6 *Upper bound on the risk-neutral cost under the true distribution against the number of neutralised attractors where the initial state is set to 70*

In the simulations, the first 64 entries of $\mathcal{R}_k$ are zeros and the last 64 are set to 10. Similarly, the first 64 entries of $\mathcal{B}_k$ are zeros and the last 64 are set to 5. The value of p_t is 0.9.

In the graphs, the risk-sensitive parameter μ takes values 0.67, 0.7 and 0.75, where according to Theorem 3 the controller with $\mu = 0.67$ is closer towards risk-neutrality or low robustness ($\mu \to 0$, which is equivalent to $s \to \infty$). And indeed, the results show, larger the μ the more robust the performance of the controller in terms of the optimal cost variation. While for initial state 0 and $\mu = 0.67$, the

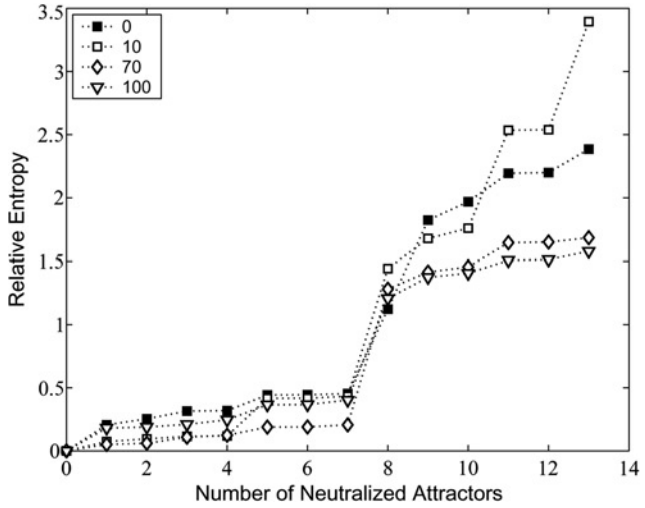

Figure 7 *Relative entropy between the nominal and true distribution against the number of neutralised attractors where the initial state is a parameter*

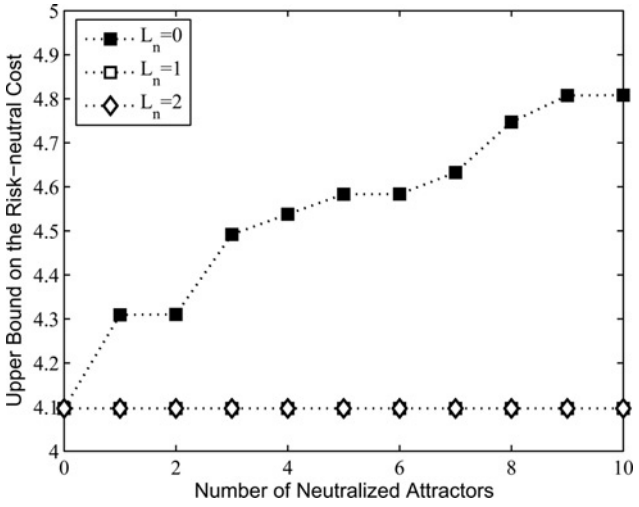

Figure 8 *Upper bound on the risk-neutral cost under the true distribution against the number of neutralised attractors and higher level states where the initial state is set to 0*

cost changes from 2.49 to 3.31 (33% change), for $\mu = 0.7$ from 4.1 to 4.88 (19% change), for $\mu = 0.75$ from 6.49 to 7.22 (11% change) with the increased distance between the nominal and the true distribution. This trend is the same regardless of the initial state of the CSPBN as seen from Figs. 5 and 6. The initial state 10 shows the most variability in the cost: for $\mu = 0.67$, the cost changes from 2.48 to 5.53 (123% change), for $\mu = 0.7$ from 4.09 to 7.01 (71% change), for $\mu = 0.75$ from 6.48 to 9.21 (42% change). Therefore a larger value of μ would lead to a better robustness.

On the other hand, at least for the example of the CSPBN considered in this paper, the larger μ the higher the cost. This is a typical behaviour of the risk-sensitive controllers for smaller uncertainties. Hence, the choice of μ will depend on how important the variability of the cost is from the point of view of a particular GRN.

For CSPBNs, having larger number of attractors, it can be expected that the uncertainty will be larger. Then, because the cost function for the smaller μ has larger variability, it may intersect the cost function for the larger μ and can be dominant for the higher uncertainties. This implies that the higher values of μ are desirable for higher uncertainties.

In practice, the lack of robustness subject to uncertainty, which would manifest as an increase of the true cost, could mean that: (1) The GRN would end up in some unwanted state, or (2) The medication is used excessively, both undesirable outcomes.

Another interesting observation is that it seems that not all attractors affect equally the instability of the cost. For instance, Fig. 7 shows that in the case of the initial state 0 the change of probabilities of attractors 65 and 72 affects the instability of the optimal cost the most. In the case of the initial state 10, these are the attractors 65, 79 and 104, and in the case of the initial state 70 this is the attractor 65. Therefore this theoretical method can be used for differentiating among attractors with respect to their influence on the controller robustness.

It is worth mentioning that the risk-sensitive cost does not depend on the uncertainty. It only depends on the nominal distribution as it can be seen from (54).

Finally, we address the effect of the uncertainty on the upper bound (54) when the probabilities of non-attractor states are perturbed. To carry out the analysis, we use a structure of four constituent deterministic BNs comprising CSPBN. Each non-attractor state belongs to a level L_n, $1 \leq L_n \leq L_{n\max}$, where the states belonging to lower levels are closer to attractors. In fact, L_n denotes the number of transitions needed to pass from any state at level L_n to its corresponding attractor. Fig. 8, shows how the upper bound on the true cost (54) changes as the function of the number of neutralised non-attractor states for a fixed L_n, $\mu = 0.7$, and initial state 0. Fig. 8 also contains the upper bound for the level $L_n = 0$, that is, the attractor states for the purpose of comparison. The figure illustrates that the perturbation of non-attractor states does not affect the performance of the controller at all. Consequently, in this particular case the robustness of the controller completely depends on the uncertainty determined by the attractor uncertainty.

6 Conclusions

One of the most serious problems in the external control of GRNs is its fragility due to the impreciseness of underlying mathematical models. The model impreciseness comes from the imperfection of measurement devices, necessary simplicity of the mathematical models and the nature of the cell

doi: 10.1049/iet-syb.2008.0121

mechanisms. To combat previously described uncertainty in the control of GRNs, we propose the risk-sensitive control. It is well known that this approach guarantees the robustness in the presence of the uncertainty with respect to the nominal model distribution. We derived the risk-sensitive controller for the CSPBNs based on the equivalence between the CSPBN and the hidden Markov model. Most importantly, we provide the assessment of the robust controller performance by using the relation between the free energy and the relative entropy. We show that the lack of knowledge in the distribution of the CSPBN attractors may lead to instability of the achieved cost function for the risk-neutral controllers in the presence of uncertainty. This may cause the GRN to end up in an unwanted state or excessive use of medications. On the other hand, the risk-sensitive controllers exhibit robustness, where the attention has to be paid to the choice of the optimal risk-sensitive parameter. Finally, the relation between the free energy and the relative entropy proved to be a good theoretical tool for distinguishing between the attractors regarding their influence on the robustness of the optimal controllers. As for the future work, we intend to address the problem of the dynamic programming algorithm complexity for the risk-sensitive optimal control. Namely, from the tree diagrams shown in Figs. 2 and 3, it can be seen that the number of possible information-states grows exponentially with the finite horizon. In addition, the number of CSPBN states grows exponentially with the number of network genes [12]. Both issues represent a computational burden limiting the applicability of algorithms for larger GRNs.

7 Acknowledgment

This work is funded by a generous support from the NSF under grant CCF-0634969.

8 References

[1] DATTA A., PAL R., CHOUDHARY A., DOUGHERTY E.R.: 'Control approaches for probabilistic gene regulatory networks', *IEEE Signal Process. Mag.*, 2007, **24**, (1), pp. 54–63

[2] FARYABI B., DATTA A., DOUGHERTY E.R.: 'On approximate stochastic control in genetic regulatory networks', *IET Syst. Biol.*, 2007, **6**, (1), pp. 361–368

[3] HUANG S.: 'Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery', *Mol. Med.*, 1999, **77**, pp. 469–480

[4] SHMULEVICH I., DOUGHERTY R.E., KIM S., ZHANG W.: 'Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks', *Bioinformatics*, 2002, **18**, pp. 261–274

[5] KAUFFMAN S.A.: 'The origins of order: self-organization and selection in evolution' (Oxford University Press, New York, 1993)

[6] SHMULEVICH I., DOUGHERTY E.R., ZHANG W.: 'Gene perturbation and intervention in probabilistic Boolean networks', *Bioinformatics*, 2002, **18**, (10), pp. 1319–1331

[7] SHMULEVICH I., DOUGHERTY E.R., ZHANG W.: 'Control of stationary behavior in probabilistic Boolean networks by means of structural intervention', *Biol. Syst.*, 2002, **10**, (2), pp. 431–446

[8] DATTA A., CHOUDHARY A., BITTNER M.L., DOUGHERTY E.R.: 'External control in markovian genetic regulatory networks', *Mach. Learn.*, 2003, **52**, (1–2), pp. 169–191

[9] DATTA A., CHOUDHARY A., BITTNER M.L., DOUGHERTY E.R.: 'External control in markovian genetic regulatory networks: the imperfect information case', *Bioinformatics*, 2004, **20**, (6), pp. 924–930

[10] CHING W.K., ZHANG S.Q., JIAO Y., AKUTSU T., WONG A.S.: 'Optimal finite-horizon control for probabilistic Boolean networks with hard constraints'. Int. Symp. Optimization and Systems Biology, 2007

[11] PAL R., DATTA A., BITTNER M.L., DOUGHERTY E.R.: 'Intervention in context-sensitive probabilistic Boolean networks', *Bioinformatics*, 2005, **21**, (7), pp. 1211–1218

[12] PAL R., DATTA A., DOUGHERTY E.R.: 'Optimal infinite horizon control for probabilistic Boolean networks', *IEEE Trans. Signal Process.*, 2006, **54**, (6), pp. 2375–2387

[13] BALAGURUNATHAN Y., WANG N., DOUGHERTY E.R., *ET AL.*: 'Noise factor analysis for cDNA microarrays', *J. Biomed. Opt.*, 2004, **4**, (9), pp. 663–678

[14] SONTAG E.D.: 'Molecular systems biology and control', *Eur. J. Control*, 2005, **11**, pp. 396–435

[15] CHARALAMBOUS C.D., DEY S., ELLIOT R.J.: 'New finitedimensional risk-sensitive filters: small-noise limits', *IEEE Trans. Autom. Control*, 1998, **43**, (10), pp. 1424–1429

[16] CHARALAMBOUS C.D.: 'The role of information state and adjoint in relating nonlinear output feedback risk-sensitive control and dynamic games', *IEEE Trans. Autom. Control*, 1997, **42**, (8), pp. 1163–1170

[17] PAL R., DATTA A., DOUGHERTY E.R.: 'Robust intervention in probabilistic Boolean networks', *IEEE Trans. Signal Process*, 2008, **56**, (3), pp. 1280–1294

[18] ELLIOTT R.J., AGGOUN L., MOORE J.B.: 'Hidden Markov models: estimation and control' (Springer-Verlang, Berlin, Germany, 1995)

[19] FERÑANDEZ-GAUCHERAND E., MARCUS S.I.: 'Risk-sensitive optimal control of hidden Markov models: structural results', *IEEE Trans. Autom. Control*, 1997, **42**, (10), pp. 1418–1422

[20] BOEL R.K., JAMES M.R., PETERSEN I.R.: 'Robustness and risk-sensitive filtering', *IEEE Trans. Autom. Control*, 1998, **43**, (10), pp. 1424–1429

[21] BITTNER M., MELTZER P., CHEN Y., *ET AL.*: 'Molecular classification of cutaneous malignant melanoma by gene expression profiling', *Nature*, 2000, **406**, (6795), pp. 536–540

[22] CHAVEZ M., ALBERT R., SONTAG E.D.: 'Robustness and fragility of Boolean models for genetic regulatory networks', *J. Theor. Biol.*, 2005, **235**, pp. 431–449

[23] CHARALAMBOUS C.D., REZAEI F.: 'Stochastic uncertain systems subject to relative entropy constraints: induced norms and monotonicity properties of minimax games', *IEEE Trans. Autom. Control*, 2007, **52**, (4), pp. 647–663

[24] WEERARATNA A.T., JIANG Y., HOSTETTER G., ROSENBLATT K., DURAY P., BITTNER M., TRENT J.M.: 'Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma', *Cancer Cell*, 2002, **1**, pp. 279–288

[25] KIM S., LI H., DOUGHERTY E.R., *ET AL.*: 'Can Markov chain models mimic biological regulation', *J. Biol. Syst.*, 2002, **10**, (4), pp. 337–357

[26] PAL R., IVANOV I., DATTA A., BITTNER M.L., DOUGHERTY E.R.: 'Generating Boolean networks with a prescribed attractor structure', *Bioinformatics*, 2005, **21**, pp. 4021–4025

9 Appendix

To prove Theorem 1, we need to prove a series of lemmas first. The most important result that uses the change-of-measure and is used in solving the risk-sensitive optimisation problem is the so-called conditional bayes theorem, which is presented below [18].

Theorem 4: (conditional Bayes theorem) suppose $(\Omega, \mathcal{F}, \mathcal{P})$ is a probability space, and $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-field. Suppose $\bar{\mathcal{P}}$ is another probability measure absolutely continuous with respect to $\mathcal{P}$, with Radon-Nikodym derivative $d\bar{\mathcal{P}}/d\mathcal{P} = \Lambda$, and E and $\bar{E}$ denote expectations with respect to $\mathcal{P}$ and $\bar{\mathcal{P}}$, respectively. Then, if ϕ is any $\bar{\mathcal{P}}$ integrable random variable, the conditional expectation of ϕ with respect to the probability measure $\bar{\mathcal{P}}$ given the sub-σ-field $\mathcal{G}$ can be constructed by

$$\bar{E}[\phi|\mathcal{G}](\omega) = \begin{cases} \dfrac{E[\Lambda\phi|\mathcal{G}](\omega)}{E[\Lambda|\mathcal{G}](\omega)} & \omega \in \{\omega \in \Omega : E[\Lambda|\mathcal{G}](\omega) > 0\} \\ 0 & \text{otherwise} \end{cases} \tag{59}$$

Here Ω denotes the universal sample space.

Proof: The proof of the theorem can be found in [18].

This theorem enables one to compute the conditional expectation under the newly introduced measure $\bar{\mathcal{P}}$ in terms of the old probability measure $\mathcal{P}$ and vice versa. In the construction of the conditional expectation under $\bar{\mathcal{P}}$, the random variable Λ is defined, which represents the change of measure. An appropriate definition of Λ is the main problem in application of the change-of-measure technique. Next, we show how Λ can be defined for our problem.

The change-of-measure is going to be applied to both observation $\boldsymbol{Y}$ and the controlled state process $\boldsymbol{X}^u$. Since we deal with the controlled state process, it is assumed that all probability distributions used are conditioned on the control $\boldsymbol{u}$. To include this fact define new complete filtrations $\{\mathcal{G}_k\}$, $\{\mathcal{F}_k\}$ and $\{\mathcal{Y}_k\}$ generated by

$$(\boldsymbol{X}_0^u, \boldsymbol{X}_1^u, \ldots, \boldsymbol{X}_k^u, \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_k, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \tag{60}$$

$$(\boldsymbol{X}_0^u, \boldsymbol{X}_1^u, \ldots, \boldsymbol{X}_k^u, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \tag{61}$$

$$(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_k, \boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \tag{62}$$

respectively. Also for convenience of derivations, we assume that the cardinalities of the sets S_X and S_Y are the same and equal to N.

First, the change of measure is carried out for the observation process $\boldsymbol{Y}$. To construct a new measure $\bar{\mathcal{P}}$ from $\mathcal{P}^u$, under which $\{\boldsymbol{Y}_k\}$ is going to be i.i.d. with $\boldsymbol{Y}_k$ having a uniform distribution over S_Y, introduce

$$\frac{d\bar{\mathcal{P}}}{d\mathcal{P}^u} = \Lambda_k \triangleq \prod_{l=1}^{k} \lambda_l \tag{63}$$

where

$$\lambda_l \triangleq \prod_{i=1}^{N} \left(\frac{N^{-1}}{c_l^i}\right)^{Y_l^i} \tag{64}$$

The inverse change of measure for the observation process is defined by

$$\frac{d\mathcal{P}^u}{d\bar{\mathcal{P}}} = \bar{\Lambda}_k \triangleq \prod_{l=1}^{k} \bar{\lambda}_l \tag{65}$$

where

$$\bar{\lambda}_l \triangleq \prod_{i=1}^{N} (Nc_l^i)^{Y_l^i} = NPr\{\boldsymbol{Y}_l | \boldsymbol{X}_{l-1}^u\} \tag{66}$$

In the next two lemmas, we prove that $\{\boldsymbol{Y}_k\}$ is i.i.d. under $\bar{\mathcal{P}}$ where $\boldsymbol{Y}_k$ has a uniform distribution over S_Y.

Lemma 2:

$$E[\lambda_{k+1}|\mathcal{G}_k] = 1 \tag{67}$$

doi: 10.1049/iet-syb.2008.0121

Proof:

$$E[\lambda_{k+1}|\mathcal{G}_k] = E\left[\prod_{i=1}^{N}\left(\frac{N^{-1}}{c_{k+1}^i}\right)^{Y_{k+1}^i}|\mathcal{G}_k\right] \tag{68}$$

$$= E\left[\sum_{i=1}^{N}\frac{1}{Nc_{k+1}^i}Y_{k+1}^i|\mathcal{G}_k\right] \tag{69}$$

$$= \sum_{i=1}^{N}\frac{1}{Nc_{k+1}^i}E[Y_{k+1}^i|\mathcal{G}_k] \tag{70}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{c_{k+1}^i}c_{k+1}^i = 1 \tag{71}$$

The equality (69) follows from the fact that $\boldsymbol{Y}_{k+1} \in S_Y$ meaning that only one entry of the vector $\boldsymbol{Y}_{k+1}$ is non-zero (equal to 1). The equality (70) follows because c_{k+1}^i (given by (26)) is measurable with respect to $\mathcal{G}_k$. □

Lemma 3: Under $\bar{\mathcal{P}}$, $\{\boldsymbol{Y}_k\}$ are i.i.d. having uniform distribution

$$\Pr\{\boldsymbol{Y}_k = \boldsymbol{e}_i\} = \frac{1}{N} \tag{72}$$

and are independent of $\boldsymbol{X}$.

Proof:

$$\bar{\Pr}\{Y_{k+1}^j = 1|\mathcal{G}_k\} = \bar{E}[\langle \boldsymbol{Y}_{k+1}, \boldsymbol{e}_j\rangle|\mathcal{G}_k] \tag{73}$$

$$= \frac{E[\Lambda_{k+1}\langle \boldsymbol{Y}_{k+1}, \boldsymbol{e}_j\rangle|\mathcal{G}_k]}{E[\Lambda_{k+1}|\mathcal{G}_k]} \tag{74}$$

$$= \frac{E[\lambda_{k+1}\langle \boldsymbol{Y}_{k+1}, \boldsymbol{e}_j\rangle|\mathcal{G}_k]}{E[\lambda_{k+1}|\mathcal{G}_k]} \tag{75}$$

$$= E[\lambda_{k+1}\langle \boldsymbol{Y}_{k+1}, \boldsymbol{e}_j\rangle|\mathcal{G}_k] \tag{76}$$

$$= E\left[\prod_{i=1}^{N}\left(\frac{N^{-1}}{c_{k+1}^i}\right)^{Y_{k+1}^i}\langle \boldsymbol{Y}_{k+1}, \boldsymbol{e}_j\rangle|\mathcal{G}_k\right] \tag{77}$$

$$= E\left[\sum_{i=1}^{N}\frac{1}{Nc_{k+1}^i}Y_{k+1}^iY_{k+1}^j|\mathcal{G}_k\right] \tag{78}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{c_{k+1}^i}E[Y_{k+1}^iY_{k+1}^j|\mathcal{G}_k] \tag{79}$$

$$= \frac{1}{Nc_{k+1}^j}E[Y_{k+1}^j|\mathcal{G}_k] \tag{80}$$

$$= \frac{1}{N} = \Pr\{Y_{k+1}^j = 1\} \tag{81}$$

The equality (74) follows from Theorem 4. The equality (75) is true since $\Lambda_{k+1} = \lambda_{k+1}\Lambda_k$ and Λ_k defined by (63) is measurable with respect to $\mathcal{G}_k$. The equality (76) is a consequence of Lemma 2. The equality (77) follows from the definition of λ_{k+1}. The equalities (78) and (80) are implied by the fact that the vector $\boldsymbol{Y}_{k+1}$ has only one non-zero entry equal to 1. The equality (79) is true because c_{k+1}^i is measurable with respect to $\mathcal{G}_k$. The equality (81) follows from the definition of c_{k+1}^j. □

Since, under $\bar{\mathcal{P}}$, $\boldsymbol{Y}$ is independent of $\boldsymbol{X^u}$, $\boldsymbol{X^u}$ is still defined by (28).

Second, the change of measure is carried out for the state process $\boldsymbol{X}^u$. The reason for this is that although $\bar{\mathcal{P}}$ is an ideal measure for $\boldsymbol{Y}$, it is not for $\boldsymbol{X}^u$. A new measure $\hat{\mathcal{P}}$ is constructed from $\bar{\mathcal{P}}$, under which $\{\boldsymbol{X}_k^u\}$ are going to be i.i.d., by introducing

$$\frac{\mathrm{d}\hat{\mathcal{P}}}{\mathrm{d}\bar{\mathcal{P}}} = \Gamma_k \triangleq \prod_{l=1}^{k}\gamma_l \tag{82}$$

where

$$\gamma_l \triangleq \prod_{i=1}^{N}\left(\frac{N^{-1}}{a_l^i}\right)^{X_l^{u,i}} \tag{83}$$

The inverse change of measure for the state process is defined by

$$\frac{\mathrm{d}\bar{\mathcal{P}}}{\mathrm{d}\hat{\mathcal{P}}} = \bar{\Gamma}_k \triangleq \prod_{l=1}^{k}\bar{\gamma}_l \tag{84}$$

where

$$\bar{\gamma}_l \triangleq \prod_{i=1}^{N}(Na_l^i)^{X_l^{u,i}} = N\Pr\{\boldsymbol{X_l^u}|\boldsymbol{X_{l-1}^u}\} \tag{85}$$

Next two lemmas prove that $\{\boldsymbol{X}_k^u\}$ are i.i.d. under $\hat{\mathcal{P}}$.

Lemma 4:

$$\bar{E}[\gamma_{k+1}|\mathcal{G}_k] = 1 \tag{86}$$

Proof: The proof is similar to the proof of Lemma 2.

Lemma 5: Under $\hat{\mathcal{P}}$, $\{\boldsymbol{X}_k^u\}$ are i.i.d. having uniform distribution

$$\hat{\Pr}\{\boldsymbol{X}_k^u = \boldsymbol{e}_i\} = \frac{1}{N} \tag{87}$$

Proof: The proof is similar to the proof of Lemma 3.

The next lemma is necessary in proving the recursive formula for the information-state.

Lemma 6:

$$\hat{E}[\bar{\Gamma}_k|\mathcal{Y}_k] = 1 \tag{88}$$

Proof:

$$\hat{E}[\bar{\Gamma}_k|\mathcal{Y}_k] = \hat{E}\left[\prod_{l=1}^{k}\prod_{i=1}^{N}(Na_l^i)^{X_l^{u,i}}|\mathcal{Y}_k\right] \tag{89}$$

$$= \hat{E}\left[\prod_{l=1}^{k}\prod_{i=1}^{N}(Na_l^i)^{X_l^{u,i}}\right] \tag{90}$$

$$= \prod_{l=1}^{k}\hat{E}\left[\prod_{i=1}^{N}(Na_l^i)^{X_l^{u,i}}\right] \tag{91}$$

$$= \prod_{l=1}^{k}\hat{E}\left[\sum_{i=1}^{N}Na_l^iX_l^{u,i}\right] \tag{92}$$

$$= \prod_{l=1}^{k}\hat{E}\left[\hat{E}\left[\sum_{i=1}^{N}Na_l^iX_l^{u,i}|\boldsymbol{X}_{l-1}^u\right]\right] \tag{93}$$

$$= \prod_{l=1}^{k}\hat{E}\left[\sum_{i=1}^{N}Na_l^i\hat{E}[X_l^{u,i}]\right] \tag{94}$$

$$= \prod_{l=1}^{k}\hat{E}\left[\sum_{i=1}^{N}a_l^i\right] \tag{95}$$

$$= 1 \tag{96}$$

The equality (90) is true because $\{\boldsymbol{X}_k^u\}$ and $\{\boldsymbol{Y}_k\}$ are independent under $\hat{\mathcal{P}}$. The equality (91) follows because $\{\boldsymbol{X}_k^u\}$ are i.i.d. under $\hat{\mathcal{P}}$. The equality (92) is implied by the fact that the vector $\boldsymbol{X}_{k+1}^u$ has only one non-zero entry equal to 1. The equality (93) is obtained by reconditioning the expectation in (92) with respect to $\boldsymbol{X}_{l-1}^u$. The equality (94) is true because a_l^i is measurable with respect to $\boldsymbol{X}_{l-1}$, and $\boldsymbol{X}_l$ is independent of $\boldsymbol{X}_{l-1}$ under $\hat{\mathcal{P}}$. The equality (95) follows because $\hat{E}[X_l^{u,i}] = \hat{E}[\langle \boldsymbol{X}_l^u, \boldsymbol{e}_i\rangle] = \langle \hat{E}[\boldsymbol{X}_l^u], \boldsymbol{e}_i\rangle = \langle 1/N, \boldsymbol{e}_i\rangle = 1/N$. The equality (96) is implied by $\sum_{i=1}^{N} a_l^i = 1$.

Proof of Theorem 1: Here $I_j(\cdot)$ represents the indicator function which is equal to one when the argument is equal to j, and zero otherwise.

$$q_{k+1}^{\mu}(j) = \langle \boldsymbol{q}_{k+1}^{\mu}, \boldsymbol{e}_j\rangle = \sum_{s=1}^{N}\boldsymbol{q}_{k+1}^{\mu}(s)I_j(s) \tag{97}$$

$$= \bar{E}\left[I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\bar{\Lambda}_{k+1}\exp\left(\mu\sum_{l=0}^{k}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right)|\mathcal{Y}_{k+1}\right] \tag{98}$$

$$= \frac{\hat{E}[I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\bar{\Lambda}_{k+1}\bar{\Gamma}_{k+1}\exp(\mu\sum_{l=0}^{k}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l))|\mathcal{Y}_{k+1}]}{\hat{E}[\bar{\Gamma}_{k+1}|\mathcal{Y}_{k+1}]} \tag{99}$$

$$= \hat{E}\left[I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\bar{\Lambda}_{k+1}\bar{\Gamma}_{k+1}\exp\left(\mu\sum_{l=0}^{k}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right)\middle|\mathcal{Y}_{k+1}\right] \tag{100}$$

$$= \hat{E}[I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j) \times \hat{E}\left[\bar{\Lambda}_k\bar{\lambda}_{k+1}\bar{\Gamma}_k\bar{\gamma}_{k+1}\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right) \times \exp(\mu L(\boldsymbol{X}_k^u, \boldsymbol{u}_k))|\boldsymbol{X}_{k+1}^u, \mathcal{Y}_{k+1}\right]\middle|\mathcal{Y}_{k+1}\right] \tag{101}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\hat{E}\left[\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}N\right)\bar{\Gamma}_k\bar{\lambda}_{k+1} \times \bar{\Lambda}_k\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right)\exp(\mu L(\boldsymbol{X}_k^u, \boldsymbol{u}_k))|\boldsymbol{X}_{k+1}^u, \mathcal{Y}_{k+1}\right] \times \hat{\Pr}\{\boldsymbol{X}_{k+1}^u|\mathcal{Y}_{k+1}\} \tag{102}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\hat{E}\left[\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}N\right)\bar{\Gamma}_k\bar{\lambda}_{k+1} \times \bar{\Lambda}_k\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right)\exp(\mu L(\boldsymbol{X}_k^u, \boldsymbol{u}_k))|\boldsymbol{X}_{k+1}^u, \mathcal{Y}_{k+1}\right] \times \frac{1}{N} \tag{103}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\hat{E}\left[\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}\right)\bar{\lambda}_{k+1}\exp(\mu L(\boldsymbol{X}_k^u, \boldsymbol{u}_k)) \times \hat{E}\left[\bar{\Gamma}_k\bar{\Lambda}_k\exp\left(\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right)\middle|\boldsymbol{X}_k^u, \boldsymbol{X}_{k+1}^u, \mathcal{Y}_{k+1}\right]\middle|\boldsymbol{X}_{k+1}^u, \mathcal{Y}_{k+1}\right] \tag{104}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\hat{E}\left[\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}\right)\bar{\lambda}_{k+1}\exp(\mu L(\boldsymbol{X}_k^u, \boldsymbol{u}_k)) \times \hat{E}\left[\bar{\Gamma}_k\bar{\Lambda}_k\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{X}_l^u, \boldsymbol{u}_l)\right)\middle|\boldsymbol{X}_k^u, \mathcal{Y}_k\right]|\boldsymbol{X}_{k+1}^u, \mathcal{Y}_{k+1}\right] \tag{105}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\frac{1}{N}\sum_{p=1}^{N}\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}\right)\bar{\lambda}_{k+1}\exp(\mu L(\boldsymbol{e}_p, \boldsymbol{u}_k)) \times \hat{E}\left[\bar{\Gamma}_k\bar{\Lambda}_k\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{e}_p, \boldsymbol{u}_l)\right)\middle|\boldsymbol{X}_k^u = \boldsymbol{e}_p, \mathcal{Y}_k\right] \tag{106}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\frac{1}{N}\sum_{p=1}^{N}\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}\right)\bar{\lambda}_{k+1}\exp(\mu L(\boldsymbol{e}_p, \boldsymbol{u}_k)) \times \hat{E}\left[I(\boldsymbol{X}_k^u = \boldsymbol{e}_p)\bar{\Gamma}_k\bar{\Lambda}_k\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{e}_p, \boldsymbol{u}_l)\right)\middle|\mathcal{Y}_k\right] \tag{107}$$

$$= I(\boldsymbol{X}_{k+1}^u = \boldsymbol{e}_j)\frac{1}{N}\sum_{p=1}^{N}\left(\sum_{i=1}^{N}a_{k+1}^iX_{k+1}^{u,i}\right)\bar{\lambda}_{k+1}\exp(\mu L(\boldsymbol{e}_p, \boldsymbol{u}_k)) \times \bar{E}\left[I(\boldsymbol{X}_k^u = \boldsymbol{e}_p)\bar{\Lambda}_k\exp\left(\mu\sum_{l=0}^{k-1}L(\boldsymbol{e}_p, \boldsymbol{u}_l)\right)\middle|\mathcal{Y}_k\right] \tag{108}$$

IET Syst. Biol., 2009, Vol. 3, Iss. 4, pp. 279–295
doi: 10.1049/iet-syb.2008.0121

$$= I(\boldsymbol{X}^u_{k+1} = \boldsymbol{e}_j)\frac{1}{N}\sum_{p=1}^{N}\left(\sum_{i=1}^{N} a^i_{k+1}X^{u,i}_{k+1}\right)\left(\sum_{r=1}^{N} c^r_{k+1}Y^{u,r}_{k+1}N\right)$$
$$\times \exp(\mu L(\boldsymbol{e}_p, \boldsymbol{u}_k))\bar{E}\left[I(\boldsymbol{X}^u_k = \boldsymbol{e}_p)\bar{\Lambda}_k.\right.$$
$$\left.\times \exp\left(\mu\sum_{l=0}^{k-1} L(\boldsymbol{e}_p, \boldsymbol{u}_l)\right)\middle|\mathcal{Y}_k\right] \quad (109)$$

$$= I(\boldsymbol{X}^u_{k+1} = \boldsymbol{e}_j)\sum_{p=1}^{N}\left(\sum_{i=1}^{N} a^i_{k+1}X^{u,i}_{k+1}\right)\left(\sum_{r=1}^{N} c^r_{k+1}Y^{u,r}_{k+1}\right)$$
$$\times \exp(\mu L(\boldsymbol{e}_p, \boldsymbol{u}_k))\boldsymbol{q}^{\mu}_k(p) \quad (110)$$

$$= \sum_{p=1}^{N}\Pr\{\boldsymbol{X}^u_{k+1} = \boldsymbol{e}_j|\boldsymbol{X}^u_k = \boldsymbol{e}_p\}\Pr\{\boldsymbol{Y}^u_{k+1}|\boldsymbol{X}^u_k = \boldsymbol{e}_p\}$$
$$\times \exp(\mu L(\boldsymbol{e}_p, \boldsymbol{u}_k))\boldsymbol{q}^{\mu}_k(p) \quad (111)$$

The equality (99) is the consequence of Theorem 4. The equality (100) follows from Lemma 6. In (101), the expectation is reconditioned on $\boldsymbol{X}^u_{k+1}$. The outer expectation in (101) is with respect to $\boldsymbol{X}^u_{k+1}$ giving (102) and (103) because, under $\hat{P}$, $\boldsymbol{X}^u_{k+1}$ is independent on $\mathcal{Y}_{k+1}$ and is uniformly distributed. In (104), the expectation is reconditioned on $\boldsymbol{X}^u_k$, and two terms are factored out from the inner expectation because c^i_{k+1} and λ_{k+1} are $\boldsymbol{X}^u_k$ measurable. In (105), the inner expectation conditioning is reduced to $\boldsymbol{X}^u_k$ and $\mathcal{Y}_k$ since the expectation argument is independent on $\boldsymbol{X}^u_{k+1}$ and $\boldsymbol{Y}_{k+1}$. The equality (106) follows because the outer expectation in (105) is the expectation with respect to $\boldsymbol{X}^u_k$ which, under $\hat{\mathcal{P}}$, is independent of $\boldsymbol{X}^u_{k+1}$ and $\mathcal{Y}_{k+1}$, and is uniformly distributed. The equality (108) is the consequence of the change-of-measures. The equality (109) is obtained by substituting the definition of λ_{k+1}. The equality (111) comes from substitution of (66) and (85) into (110).