Context-Sensitive Probabilistic Boolean Networks: Steady-State Properties, Reduction, and Steady-State Approximation

Ranadip Pal, *Member, IEEE*

***Abstract*—Context-sensitive probabilistic Boolean networks (PBNs) have been recently introduced as a paradigm for modeling genetic regulatory networks and have served as the main model for the application of intervention methods, including optimal control strategies, to favorably effect system dynamics. Since it is believed that the steady-state behavior of a context-sensitive PBN is indicative of the phenotype, it is important to study the alternation in the steady-state probability distribution due to any variations in the formulations of the context-sensitive PBNs. Furthermore, the huge computational complexity of the context-sensitive PBN model necessitates generation of size-reduction techniques and approximate methods for calculation of the steady-state probability distribution of context-sensitive PBNs. The goal of this paper is threefold: i) to study the effects of the various definitions of context-sensitive PBNs on the steady-state probability distributions and the downstream control policy design; ii) to propose a reduction technique that maintains the steady-state probability distribution; and iii) to provide an approximation method for calculating the steady-state probability distribution of a context-sensitive PBN.**

***Index Terms*—Context-sensitive PBN, genetic regulatory network model, steady state properties, reduction mapping, steady state distribution approximation.**

I. Introduction

GIVEN a set of genes, the evolution of their expression levels constitutes a dynamical system over time. A large number of approaches have been proposed to model the behavior of gene regulatory networks, both deterministic and stochastic [1], [2]. Some models use biochemical information to capture fine details involving transcription and translation via differential equations or their discrete approximations [3]–[6]. Others use data to infer the parameters of differential or difference equation models [7]–[11]. Discrete models with synchronous timing have been used since the late 1960s in the form of the Boolean network (BN) model [12]–[14]. In this model, gene expression is quantized to two levels: ON and OFF, denoted by 1 and 0, respectively. The expression level (state) of a gene is functionally related via a logical rule to the expression states of other genes. Specifically, a BN is composed of a set $V = \{x_1, x_2, \ldots, x_n\}$ consisting of n binary variables, each denoting a gene expression, and a set $F = \{f_1, f_2, \ldots, f_n\}$ of regulatory functions, such that for discrete time, $t = 0, 1, 2, \ldots, x_i(t+1) = f_i(x_1(t), \ldots, x_n(t))$. At any time point, the state of the network is given by an expression vector $(x_1(t), \ldots, x_n(t))$, called the gene activity profile (GAP). The BN model has yielded insights into the overall behavior of large genetic networks [15]–[18]. The dynamic behavior of BNs can be used to model many biologically meaningful phenomena—for instance, cellular state dynamics possessing switch-like behavior, stability, and hysteresis [19]. There has also been a number of attempts to model the cell cycle using Boolean networks [20]–[22]. Besides the conceptual framework afforded by such models, practical uses, such as the identification of suitable drug targets in cancer therapy, result by inferring the structure of the genetic models from experimental data, e.g., from the gene expression profiles [19].

In the Boolean model, the assumption of a single transition rule for each gene can be problematic with respect to inference. The data are typically noisy, the number of samples is small relative to the number of parameters to be estimated, and inputs to the system are typically not controlled. Thus, one may observe various possible transitions from a given state. A salient instance is that unobserved (latent) variables external to the model network result in the network structure taking on various realizations. Owing to these considerations, a stochastic rule-based regulatory model, the probabilistic Boolean network (PBN) [23], [24] has been proposed. A PBN is composed of a collection $\{B_1, B_2, \ldots, B_k\}$ of finite-state-space networks, each called a *context* and each composed of the same set of genes. At each time point, a random decision is made whether to switch (with probability q) to a new context or remain (with probability $1-q$) in the current context. If a switch is made, the choice of context is governed by a probability density $\{c_1, \ldots, c_k\}$. If $q = 1$, meaning certain switching, then the PBN is said to be *instantaneously random;* if $q < 1$, with a small switching probability meant to capture the effect of latent variables, then the PBN is said to be *context-sensitive* [25], [26]. Limiting the switching between the constituent networks increases the stability of the biological system. In both cases, the probabilistic structure of the PBN can be modeled as a Markov chain. In the instantaneously random case, the states of the Markov chain are the GAPs; in the context-sensitive case, the states of the Markov chain consist of (context, GAP) pairs.

Manuscript received February 05, 2009; accepted July 26, 2009. First published August 21, 2009; current version published January 13, 2010. The associate editor coordinating review of this manuscript and approving it for publication was Dr. Yufei Huang.

The author is with Texas Tech University, Electrical and Computer Engineering, Lubbock, TX, 79409 USA (e-mail: ranadip.pal@ttu.edu).

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

Digital Object Identifier 10.1109/TSP.2009.2030832

If we assume that, at any time point and for any gene, there is a positive probability (p) of random perturbation (owing to some event outside the network model), then the associated Markov chain possesses a steady-state distribution [25]. The rule-based structure of PBNs mirrors the functional transcriptional regulation in cells, and their stochastic structure allows them to reflect uncertainty, both in the data and in model selection.

Definitions of the context-sensitive PBNs in the literature [25]–[27] differ with respect to whether at time step t, network switching takes place before or after the state transition. In this paper, we will show that this minor variation in formulation of the problem can substantially change the steady-state probabilities of the states. The generated relationships between the steady-state probability distributions of the two formulations will assist in adapting the existing state reduction and intervention algorithms designed using one of the formulations to the other formulation. We will also illustrate the effectiveness of one formulation over another in the calculation of one-step intervention policies. Furthermore, the generated relationships between the steady-state probability distributions of the two formulations will ease understanding of context-reduction mappings discussed later in the paper.

Context-reduction mappings refer to the generation of an approximate transition probability matrix, absent of context, for the Markov chain associated with the context-sensitive PBN. For instance, if we have n binary genes and k contexts, then a context-sensitive PBN denoted by Φ will consist of a $2^n k \times 2^n k$ size transition probability matrix P_Φ, while the context-reduced PBN denoted by ϕ will have a $2^n \times 2^n$ size transition probability matrix P_ϕ. The steady-state probability of a GAP state in the context-sensitive PBN Φ can be calculated by summing over the contexts. Suppose π denotes the $2^n k$ length steady-state probability vector of the context-sensitive PBN Φ with state $2^n(x-1)+i$ corresponding to BN x and GAP state i. Then the steady-state probability of GAP state i in the context-sensitive PBN is equal to $\sum_{x=1}^{k} \pi(2^n(x-1)+i)$. Since the steady-state probabilities are important characteristics of the network, our interest lies in maintaining the steady-state probabilities of the GAP states after the context reduction. For instance, if η represents the 2^n length steady-state probability vector of the context-reduced PBN ϕ, we would like to have $\eta(i) = \sum_{x=1}^{k} \pi(2^n(x-1)+i)$. Earlier approaches to context-reduction mappings [26]–[28] did not satisfy this criterion. The second goal of this paper is to provide a context-reduction mapping that maintains the steady-state probability distribution of the GAP states. A set of other approaches on reducing the number of genes in a probabilistic Boolean network has been studied in the literature [29]–[31], but none of them has proved equivalence of the steady-state probabilities of the reduced network and the collapsed steady-state probabilities of the original context-sensitive PBN.

The third and final goal of this paper is to propose an approximation algorithm for generation of the steady-state probabilities of a context-sensitive PBN. The huge computational complexity associated with the calculation of the steady-state probability distribution of a $2^n k \times 2^n k$ matrix necessitates the design of fast approximation methods. Previous approximation methods for calculation of steady-state probabilities of a PBN have considered generation of the steady-state probabilities of the major attractors of the constituent Boolean networks [25] or removal of Boolean networks with small selection probabilities [32]. In this paper, we will approach the approximation based on the steady-state probability distribution of the constituent Boolean networks. The proposed method provides good approximation for small values of network transition probability q and has a time complexity of $O(2^{3n}k)$ as compared to a time complexity of $O(2^{3n}k^3)$ for calculation of the true steady-state probabilities of the context-sensitive PBN. We will also derive analytical bounds for the errors due to the approximation.

This paper is organized as follows. Section II provides the different formulations of the context-sensitive PBNs and the relationships between them. We present the context-reduction mapping in Section III . Section IV contains an algorithm to approximate the steady-state probability distribution of the context-sensitive PBN. Conclusions are presented in Section V.

II. Formulations of Context-Sensitive PBNs

A. Transition Probabilities of Context-Sensitive PBN

We consider a context-sensitive PBN Φ consisting of n genes and composed of k Boolean networks with selection probabilities $c_1, c_2, \ldots, c_k$. Furthermore, let p and q be the perturbation probability and the network switching probability respectively. Clearly, each Boolean network will have 2^n states zero to 2^n-1 (equivalent binary representations $00\cdots0$ to $11\cdots1$) and the collection of k BNs can be considered to have $2^n k$ states. The initial definition of context-sensitive PBNs in [25] allowed four mutually exclusive events to occur at time t.

- $\Phi1$: The current network function is applied, the PBN transitions accordingly, and the network function remains the same for the next transition.
- $\Phi2$: The current network function is applied, the PBN transitions accordingly, and a new network function is selected for the next transition.
- $\Phi3$: There is a random perturbation and the network function remains the same for the next transition.
- $\Phi4$: There is a random perturbation and a new network function is selected for the next transition.

Let $w(t) \in \{0, 1, 2, \ldots, 2^n k - 1\}$ be the state that is occupied by the PBN at time t. w consists of the decimal equivalent of the binary context and the GAP information. Let $a, b \in \{0, 1, 2, \ldots, 2^n k - 1\}$ be any two states and define $r_1 = \lceil (a+1)/2^n \rceil$,[1] $r_2 = \lceil (b+1)/2^n \rceil$, $u_1 = a - 2^n(r_1 - 1)$, and $u_2 = b - 2^n(r_2 - 1)$. r_1 and r_2 denotes the BNs corresponding to states a and b, respectively, whereas u_1 and u_2 signifies the decimal representation of the n GAP of states a and b, respectively. We will assume that if at time t, the network is at BN r, then at time $t+1$, we will remain at BN r with probability $1 - q(\sum_{l=1, l\neq r}^{k} c_l)$ or transition to BN s with probability qc_s, where $s \in 0, 1, \ldots, k$, and $s \neq r$. There is a minor difference in this assumption [27], [33] and the assumption in [26] and [28], where the probability of remaining at network r at time $t+1$ was assumed to be $1-q$, i.e., returning to the same network function was not allowed for events $\Phi2$ and $\Phi4$. We are incorporating this change, as the new definition will ease the understanding of

[1] $\lceil x \rceil = \min(n \in \mathbb{Z} | n \geq x)$.

context switching for the specific case of two contexts. To illustrate that, let us consider an example of a context-sensitive PBN consisting of two contexts $\mathbf{f}_1$ and $\mathbf{f}_2$ with network selection probabilities c_1 and c_2 and network transition probability q. In the previous formulation, if we are at network $\mathbf{f}_1$ at time t, the network function changes to $\mathbf{f}_2$ at time $t+1$ with probability q irrespective of c_1 and c_2. With the new assumption, the network can change to $\mathbf{f}_2$ with probability qc_2 or remain at $\mathbf{f}_1$ with probability $1-q+qc_1$.

Then based on the new assumption and the definition in [25], the one-step transition probability from a to b is given by

$$\begin{aligned} &P_\Phi\left(w(t+1)=b|w(t)=a\right) \\ &\quad=\left\{(1-p)^n f_{r_1,u_1,u_2}+(1-p)^{n-h}p^h s(h)\right\} \\ &\qquad\times\left[(1-q+qc_{r_1})\,g(a,b)+qc_{r_2}\,(1-g(a,b))\right]. \end{aligned} \quad (1)$$

where h is the Hamming distance between u_1 and u_2, i.e., the number of genes that differ between the two states

$$f_{r,u_1,u_2}=\begin{cases}1, & \text{if } u_1 \text{ transitions to } u_2 \text{ in a} \\ & \quad \text{single step in network } r \\ 0, & \text{otherwise}\end{cases} \quad (2)$$

$$g(a,b)=\begin{cases}1, & \text{if } r_1=r_2=r \\ 0, & \text{otherwise}\end{cases} \quad (3)$$

and

$$s(h)=\begin{cases}0, & \text{if } h=0 \\ 1, & \text{otherwise.}\end{cases} \quad (4)$$

To visualize the transition probability matrix better, let us denote by P_i the $2^n \times 2^n$ transition probability matrix corresponding to BN i.[2] In other words, $P_i(z_{t+1}=u_2|z_t=u_1)=(1-p)^n f_{i,u_1,u_2}+(1-p)^{n-h}p^h s(h)$, where $z_t \in 0,1,\ldots,2^n-1$ denotes the decimal representation of the GAP state of a BN at time t and h is the Hamming distance between u_1 and u_2, $u_1,u_2 \in 0,1\cdots 2^n-1$. Then the $2^n k \times 2^n k$ transition probability matrix P_Φ corresponding to context-sensitive PBN Φ is given by

$$P_\Phi=\begin{bmatrix}(1-q+qc_1)P_1 & qc_2P_1 & \cdots & qc_kP_1 \\ \cdots & \cdots & \cdots & \cdots \\ qc_1P_k & qc_2P_k & \cdots & (1-q+qc_k)P_k\end{bmatrix}. \quad (5)$$

B. Ordering of Network Switching and State Transitions

If we revisit the definition of event $\Phi 2$ for a context-sensitive PBN, we notice that the current network function is applied first and then the network switch occurs. Based on this definition, if at time t, the network is r_1 and the GAP is u_1, then the GAP at time $t+1$ is u_2, where u_2 is dependent on r_1 alone and not dependent on whether a network switch occurs or not. The statement is evident when we consider the matrix in (5), where each row is dependent on a single P_i. We next show through an example that this definition is not suitable for a one-step control problem where the cost criteria are based on the GAP alone. Let us consider a context-sensitive PBN Φ_1 consisting of two BNs of three genes (eight states). If the 8×8 size transition probabilities of the BNs with perturbation are given by P_1 and P_2, then the 16×16 size transition probability matrix of the context-sensitive PBN Φ_1 is given by

$$P_{\Phi_1}=\begin{bmatrix}(1-q+qc_1)P_1 & qc_2P_1 \\ qc_1P_2 & (1-q+qc_2)P_2\end{bmatrix}. \quad (6)$$

Previous works [26], [34]–[36] on intervention in probabilistic Boolean networks had assumed that the status of one of the genes was indicative of the cost of the state and that the control action consisted of flipping one of the other genes. Let us assume those conditions and set up a *finite horizon* control problem similar to [26] and [34]. Given an initial state z_0

$$\min_{\mu_0,\mu_1,\ldots,\mu_{M-1}} E\left[\sum_{t=0}^{M-1} g_t\left(w_t,\mu_t(w_t)\right)+g_M(w_M)\right] \quad (7)$$

subject to the control-dependent one-step transition probability $p_{ij}(\varphi)=Pr(w_{t+1}=j|w_t=i,\varphi_t=\varphi)$, where we have the following:

- M represents the treatment/intervention window;
- μ_t for $t=0,1,2,\ldots,M-1$ are functions mapping the state space into the control space, i.e., the controls considered are state feedbacks; for n genes and k contexts, the state space is 0,1,…, $k2^n-1$ and for m binary controls, the control space is 0,1,…, 2^m-1;
- $g_t(w_t,\varphi_t)$ is the one-step cost of applying the control φ_t at state w_t;
- $g_M(w_M)$ is the terminal cost associated with the state z_M.

The dynamic programming solution to (7) is given by [34], [37]

$$J_M(w_M)=g_M(w_M) \quad (8)$$

$$J_t(w_t)=\min_{\varphi_t \in C}\left[g_t(w_t,\varphi_t)+\sum_{j=0}^{16-1} p_{w_t,j}(\varphi_t)J_{t+1}(j)\right] \quad t=M-1,M-2,\ldots,1,0. \quad (9)$$

For instance, state 9 refers to binary 1001 with the most significant bit 1 representing the context, the next bit with value 0 representing the state of the first gene, and so on. We observe that if the control is flipping the second gene and the cost is based on the first gene, the control policy decision at time $t=M-1$ for states 0–7 (belonging to the first context) will not be affected by P_2. To illustrate this, let us consider the minimization equation $g_{M-1}(w_{M-1},\varphi_{M-1})+\sum_{j=0}^{16-1} p_{w_{M-1},j}(\varphi_{M-1})J_M(j)$ for $w_{M-1} \in \{0,1,\ldots,7\}$. The minimization will not be dependent on P_2, as the first eight rows of the matrix P_{Φ_1} are not dependent on P_2 and control by flipping the second gene switches the transition probabilities from among the first eight rows.[3] Furthermore, for J_M dependent on GAP and independent of context, the minimization term at stage $t=M-1$ will not be dependent on q. Thus, for a single intervention step, this formulation will not consider the effect of q and the transition probabilities of the other networks.

[2]Note that P_i here refers to the transition probability matrix of a BN with perturbation.

[3]Control by flipping a gene changes the initial state of the network. For instance, consider the case of eight GAP states (three genes) and two context with MSB denoting the context. If the second gene (third binary digit) in the state 0110 is flipped, the initial state becomes 0100. Consequently, in terms of transition probabilities, the seventh row corresponding to 0110 will be replaced by the fifth row corresponding to 0100.

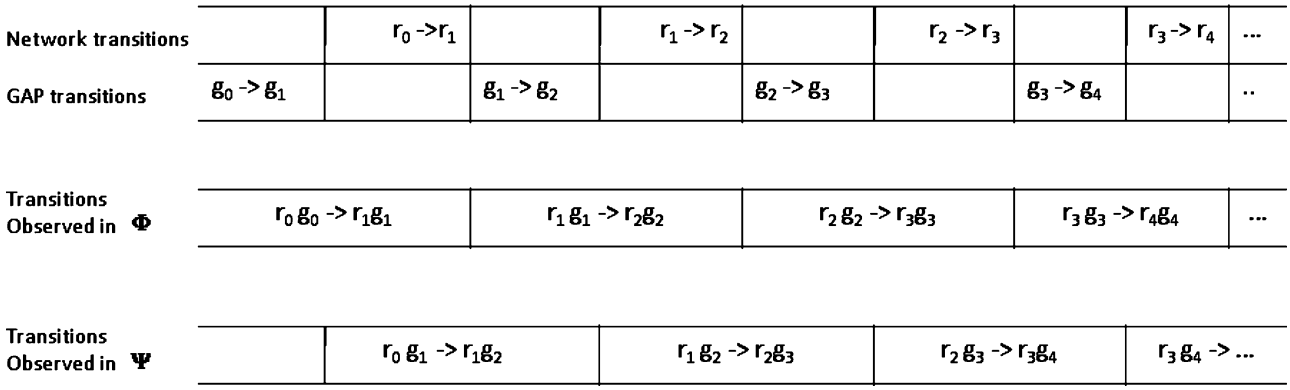

Fig. 1. Transitions for formulation Φ and Ψ.

We will henceforth consider another formulation that will incorporate the effect of q and the other network transitions in a single step. Furthermore, the new formulation and the associated relationships between the two formulations will increase our overall analytical understanding of the context switching process. The new formulation has been recently used in [27] to compare the intervention performance in case of model reduction.

Let us denote the context-sensitive PBN corresponding to the new formulation by Ψ. The four mutually exclusive events that can occur at time t for Ψ are as follows.

- $\Psi 1$: The current network function is applied, the PBN transitions accordingly, and the network function remains the same for the next transition.
- $\Psi 2$: A new network function is selected and the PBN transitions according to the new network function.
- $\Psi 3$: There is a random perturbation and the network function remains the same for the next transition.
- $\Psi 4$: A new network function is selected and there is a random perturbation.

Similar to Φ, the one-step transition probability from state a to b for Ψ is given by

$$\begin{aligned} &P_\Psi\left(w(t+1)=b|w(t)=a\right) \\ &\quad=\left[(1-q+qc_{r_1})\right) \\ &\qquad\times\left\{(1-p)^n f_{r_1,u_1,u_2}+(1-p)^{n-h}p^h s(h)\right\}\big]\, g(a,b) \\ &\qquad+\left[qc_{r_2}\left\{(1-p)^n f_{r_2,u_1,u_2}+(1-p)^{n-h}p^h s(h)\right\}\right] \\ &\qquad\times(1-g(a,b)) \end{aligned} \tag{10}$$

which is same as (1) except that f_{r_2,u_1,u_2} has been used in place of f_{r_1,u_1,u_2} for the second term.

The $2^n k \times 2^n k$ transition probability matrix corresponding to Ψ can be represented as

$$P_\Psi=\begin{bmatrix} (1-q+qc_1)P_1 & qc_2P_2 & \cdots & qc_kP_k \\ \cdots & \cdots & \cdots & \cdots \\ qc_1P_1 & qc_2P_2 & \cdots & (1-q+qc_k)P_k \end{bmatrix}. \tag{11}$$

Equation (11) shows that each row of P_Ψ is dependent on all the constituent BNs and not dependent on a single BN as in the case of P_Φ. Furthermore, if we consider the minimization term in (9), we will notice that this formulation takes into consideration the effect of q at step $t=M-1$ of the intervention problem.

Fig. 1 provides a graphical representation of the two formulations Φ and Ψ. The network and the GAP transitions of the context-sensitive PBN at consecutive time instants are shown in the first two rows, respectively. The transitions will be observed in different ways by the two formulations: Φ will consider the GAP transitions first and then the network transitions, whereas Ψ will consider the network transitions first and then the GAP transitions. The observed transitions for the two formulations are shown in the third and fourth rows of the figure. For instance, Ψ will consider the network transition $r_0 \rightarrow r_1$ first and then the GAP transition $g_1 \rightarrow g_2$, and thus we will observe the transition $r_0 g_1 \rightarrow r_1 g_2$ if formulation Ψ is used.

Let the $2^n k$ length steady-state probability vectors for Φ and Ψ be denoted by $\pi_\Phi(0)\cdots\pi_\Phi(2^n k-1)$ and $\pi_\Psi(0)\cdots\pi_\Psi(2^n k-1)$, respectively. The relationship between the two steady-state probabilities is given by the following theorem.

Theorem II.1: $\pi_\Phi(a)=(1-q)\pi_\Psi(a)+qc_r\sum_{i=1}^{k}\pi_\Psi(2^n(i-1)+a_u)$, where $r=\lceil (a+1)/2^n\rceil$ and $a_u=a-2^n(r-1)$.

Proof: Let us consider long runs of the network corresponding to formulations Φ and Ψ. When there are no network transitions, both the formulations will behave exactly the same. When there is a network transition from network i to r, the qc_r part of the steady-state probability corresponding to state $2^n(i-1)+a_u$ in formulation Ψ will be reflected in the steady-state probability of a for formulation Φ. This is because during a network transition from network i to network r, the GAP transition occurring will be according to network i in formulation Φ, and qc_r is the probability of transition from network i to network r, thus adding $qc_r\pi_\Psi(2^n(i-1)+a_u)$ to $\pi_\Phi(a)$. In other words, a proportion of GAP profiles from the last networks will turn up as GAP profiles of the new network.

For instance, during the network transition $r_1 \rightarrow r_2$ in Fig. 1, the GAP g_2 will be considered in network r_1 for formulation Ψ, while it will be considered in network r_2 for formulation Φ. Similarly, when a network transition occurs from r to another network j, the $q(1-c_r)$ part of the steady-state probability corresponding to state a in formulation Ψ will not be reflected in the steady-state probability of a for formulation Φ.

Thus

$$\begin{aligned} \pi_\Phi(a)=&\pi_\Psi(a)-q\pi_\Psi(a)(1-c_r) \\ &+qc_r\sum_{i=1,i\neq r}^{k}\pi_\Psi\left(2^n(i-1)+a_u\right) \end{aligned} \tag{12}$$

$$= \pi_\Psi(a) - q\pi_\Psi(a)(1-c_r) + \left(\sum_{i=1}^{k} \pi_\Psi\left(2^n(i-1)+a_u\right) - \pi_\Psi(a)\right) qc_r \quad (13)$$

$$= \pi_\Psi(a) - q\pi_\Psi(a) + qc_r \sum_{i=1}^{k} \pi_\Psi\left(2^n(i-1)+a_u\right)$$

$$= (1-q)\pi_\Psi(a) + qc_r \sum_{i=1}^{k} \pi_\Psi\left(2^n(i-1)+a_u\right). \quad (14)$$

The summation term in (12) is due to the fact that i can be any of the networks except r.

Corollary II.2: $\sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Phi(a) = (1-q) \sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Psi(a) + qc_r$.

The next theorem states that the sum of the steady-state probabilities of the GAP states for a particular context r is equal to the selection probability of the context c_r for both the formulations.

Theorem II.3: $\sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Phi(a) = \sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Psi(a) = c_r$.

Proof: Let us consider the original steady-state probability equations ($\pi_\Psi = \pi_\Psi P_\Psi$) for the formulation Ψ given by (11). The corresponding equations for $a = 2^n(r-1)+v$ are

$$\pi_\Psi(a) = (1-q) \sum_{u=0}^{2^n-1} \pi_\Psi\left(2^n(r-1)+u\right) P_r(u,v) + qc_r \sum_{i=1}^{k} \sum_{u=0}^{2^n-1} \pi_\Psi\left(2^n(i-1)+u\right) P_r(u,v) \quad (15)$$

where $r \in [1,\ldots,n]$ and $v \in [0,1,\ldots,2^n-1]$.

Thus

$$\begin{aligned}
&\sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Psi(a) \\
&= (1-q) \sum_{u=0}^{2^n-1} \pi_\Psi\left(2^n(r-1)+u\right) \sum_{v=0}^{2^n-1} P_r(u,v) \\
&\quad + qc_r \sum_{i=1}^{k} \sum_{u=0}^{2^n-1} \pi_\Psi\left(2^n(i-1)+u\right) \sum_{v=0}^{2^n-1} P_r(u,v) \\
&= (1-q) \sum_{u=0}^{2^n-1} \pi_\Psi\left(2^n(r-1)+u\right) \\
&\quad + qc_r \sum_{i=1}^{k} \sum_{u=0}^{2^n-1} \pi_\Psi\left(2^n(i-1)+u\right) \\
&= (1-q) \sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Psi(a) + qc_r. \quad (16)
\end{aligned}$$

Therefore, $\sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Psi(a) = c_r$.

From Corollary II.2, we arrive at $\sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Phi(a) = \sum_{a=2^n(r-1)}^{2^n r-1} \pi_\Psi(a) = c_r$.

C. Simulations

Simulations were conducted to illustrate the difference in the steady-state probabilities of the two formulations and to verify (14). We generated 100 random context-sensitive PBNs for different combinations of n, p, q, and k and noted the differences between the steady-state probability distributions of formulations Φ and Ψ. The $\|.\|_1$, $\|.\|_2$, and $\|.\|_\infty$ norms were used as the metrics for measuring the difference in the steady-state probability distributions. The range of n, p, q, and k was $n = [2,3,4,5,6,7]$, $p = [.01,.05,.1,.2,.25,.3,.4,.5,.75,.99]$, $q = [.01,.05,.1,.2,.25,.3,.4,.5,.75,1]$, and $k = [2,5,10]$, which produced a set of 1800 different combinations of n,p,q,k. To generate a set of k random c_is, a set of $k-1$ points were selected uniformly randomly between zero and one. The numbers 0 and 1 were appended to the array and the array sorted. The difference of two consecutive entries in the sorted array produced the c_is. To be concise, we will present the results of only few of the combinations in Table I.[4] The fifth and sixth columns contain the average 1-norm and ∞-norm of the difference in the steady-state probability distributions of Φ and Ψ. The average is over the 100 random context-sensitive PBNs that were generated for each combination of $[n,k,p,q]$. Similarly, the seventh and eighth columns contain the average 1-norm and ∞-norm of ϵ, where ϵ denotes the difference between the true π_Φ and the π_Φ calculated using (14). We notice that the 1-norm and ∞-norm of ϵ are in the range of 10^{-15}, which is due to the Matlab numerical approximations and, in all practicality, is equal to zero. The 1-norm and ∞-norm of $\pi_\Phi - \pi_\Psi$ are, however, significant, especially for larger q.

TABLE I
ERROR TABLE

n	k	p	q	$\|\|\pi_\Phi - \pi_\Psi\|\|_1$	$\|\|\pi_\Phi - \pi_\Psi\|\|_\infty$	$\|\|\epsilon\|\|_1$	$\|\|\epsilon\|\|_\infty$
7	5	.01	.01	9.59E-3	3.71E-4	7.20E-15	5.48E-16
7	5	.01	.1	8.57E-2	2.41E-3	1.14E-15	8.04E-17
7	5	.01	.5	3.57E-1	6.36E-3	7.00E-16	3.86E-17
7	5	.01	1	6.55E-1	8.79E-3	7.27E-16	5.05E-17
7	5	.1	.01	3.34E-3	4.44E-5	3.08E-15	4.06E-17
7	5	.5	.01	4.78E-5	5.42E-7	2.92E-15	2.20E-17

Rows 2–5 of Table I show that the difference in the steady-state probabilities of the two formulations increases with an increase in q. This trend is also observable for other values of n and k. The difference in the steady-state probability distributions of the two formulations with respect to q and n for fixed $p = 0.01$ and $k = 5$ is plotted in Fig. 2. We notice that the difference increases with an increase in the value of q. We can possibly explain that by rewriting (14) in the following form: $\pi_\Phi(a) - \pi_\Psi(a) = q(c_r \sum_{i=1}^{k} \pi_\Psi(2^n(i-1)+a_u) - \pi_\Psi(a))$. This suggest that the difference increases for large q.[5] Furthermore, when q is high, the frequency of network transitions will increase, and as we have elaborated earlier that the difference in the two formulations exists only during the network transitions, consequently a higher q will tend to produce a higher difference

[4]Detailed simulation results are available at http://cvial.ece.ttu.edu/ranadippal/cspbn/.

[5]We also have to consider the variation of the term $c_r \sum_{i=1}^{k} \pi_\Psi(2^n(i-1)+a_u) - \pi_\Psi(a)$ with change in q but, assuming that the change in that term is insignificant with changes in q, a direct proportional relationship between the difference of the steady-state probability distributions and q is noticeable.

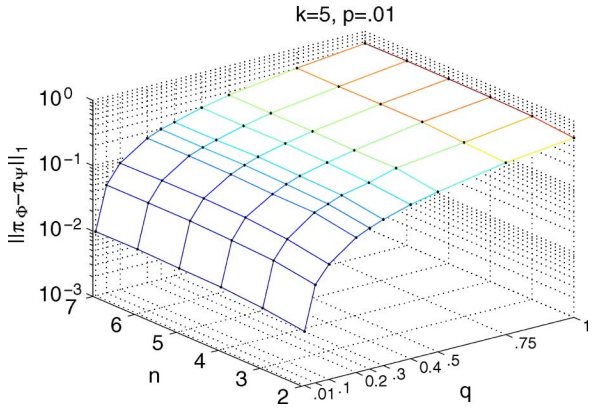

Fig. 2. Difference in the steady-state distributions of Ψ and Φ for different n and q.

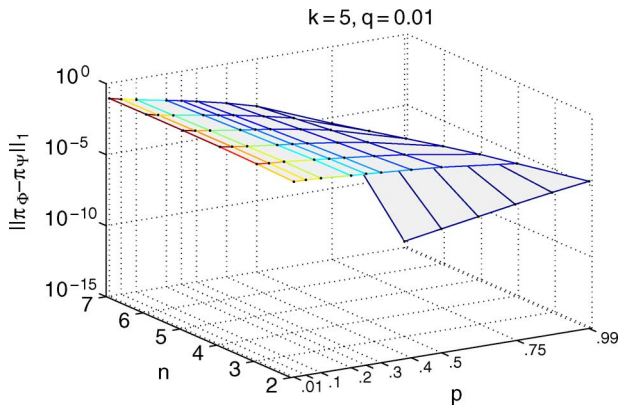

Fig. 3. Difference in the steady-state distributions of Ψ and Φ for different n and p.

in the steady-state probability distributions. We also observe that the difference in the steady-state probability distributions of the two formulations decreases with increase in p. To explain this behavior, we should note that p represents perturbation probability; and the higher the value of p, the GAP transitions will tend to follow random order as compared to following the transitions of the individual BNs of the context-sensitive PBN. Increase in randomness will make the steady-state probabilities more uniform and consequently decreasing $c_r \sum_{i=1}^{k} \pi_\Psi(2^n(i-1)+a_u) - \pi_\Psi(a)$. The variation with respect to p and n for fixed q and k is plotted in Fig. 3.

III. Steady-State Preserving Context Reduction

In this section, we will consider the collapsed steady-state probabilities, which refers to steady-state probabilities for the GAP states only. Let η_Φ and η_Ψ denote the collapsed steady-state probabilities for formulation Φ and Ψ, respectively, i.e., $\eta_\Phi(i) = \sum_{l=1}^{k} \pi_\Phi(2^n(l-1)+i)$ and $\eta_\Psi(i) = \sum_{l=1}^{k} \pi_\Psi(2^n(l-1)+i)$ for $i = 0$ to $2^n - 1$. Then

$$\begin{aligned}
\eta_\Phi(i) &= \sum_{l=1}^{k} \pi_\Phi\left(2^n(l-1)+i\right) \\
&= \sum_{l=1}^{k} \Big(\pi_\Psi\left(2^n(l-1)+i\right) - q(1-c_l)\pi_\Psi\left(2^n(l-1)+i\right) \\
&\qquad + qc_l \left(\sum_{j=1}^{k} \pi_\Psi\left(2^n(j-1)+i\right) \right. \\
&\qquad \left. - \pi_\Psi\left(2^n(l-1)+i\right) \right) \Big)
\end{aligned}$$

from ((13))

$$\begin{aligned}
&= \sum_{l=1}^{k} (1-q+qc_l)\pi_\Psi\left(2^n(l-1)+i\right) \\
&\quad + \sum_{j=1}^{k} \pi_\Psi\left(2^n(j-1)+i\right) \sum_{l=1}^{k} qc_l \\
&\quad - \sum_{l=1}^{k} qc_l \pi_\Psi\left(2^n(l-1)+i\right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{l=1}^{k} (1-q+qc_l)\pi_\Psi\left(2^n(l-1)+i\right) \\
&\quad + q \sum_{j=1}^{k} \pi_\Psi\left(2^n(j-1)+i\right) - \sum_{l=1}^{k} qc_l \pi_\Psi\left(2^n(l-1)+i\right) \\
&= \sum_{l=1}^{k} \pi_\Psi\left(2^n(l-1)+i\right) = \eta_\Psi(i). \qquad (17)
\end{aligned}$$

The above equation shows that the collapsed steady-state probabilities are not changed when we use either of the definitions Φ or Ψ. From now onwards, we will denote by $\eta = \eta_\Phi = \eta_\Psi$, the collapsed steady-state probability distribution. Thus, the relationship between π_Φ and π_Ψ provided in (14) can be simplified as shown next

$$\pi_\Phi(a) = (1-q)\pi_\Psi(a) + \eta(a_u)qc_r \qquad (18)$$

where $a = 2^n(r-1) + a_u$.

The collapsed steady-state probability distribution η is calculated from the steady-state probabilities of a network of size $2^n k \times 2^n k$. We are interested in generating a $2^n \times 2^n$ transition probability matrix that produces η as the steady-state probability distribution vector. The averaging method used in [26] to collapse the transition probability matrix will not produce steady-state probabilities equal to η. Let us revisit the problem and strengthen the simplified assumptions used in [26] to arrive at a more specific representation of the collapsed network. Here, collapsing refers to removing the context information.

Let z_t denote the GAP at time t and z_{t+1} denote the GAP at time $t+1$. The context information has been removed, and this z can take values from zero to $2^n - 1$. Similarly, let w_t ranging from zero to $2^n k - 1$ denote the context and GAP information at time t corresponding to either formulation Φ or Ψ. The transition probability of going from GAP state s_1 to GAP state s_2 is given by (19), shown at the bottom of the next page. In [26], $P(w_t = 2^n(j-1)+s_2) / \sum_{j'=1}^{k} P(w_t = 2^n(j'-1)+s_2)$ was approximated by c_j. The approximation will be very close for high values of q, as shown next. From (18), we have

$$\frac{\pi_\Phi(a)}{\eta(a_u)} = \frac{(1-q)\pi_\Psi(a)}{\eta(a_u)} + qc_r. \qquad (20)$$

Substituting $q = 1$ in (20) results in $c_j = \pi_\Phi(2^n(j-1)+s_2)/\sum_{j'=1}^{k}\pi_\Phi(2^n(j'-1)+s_2)$. For large t, $P(w_t = 2^n(j-1)+s_2)/\sum_{j'=1}^{k}P(w_t = 2^n(j'-1)+s_2) = \pi_\Phi(2^n(j-1)+s_2)/\sum_{j'=1}^{k}\pi_\Phi(2^n(j'-1)+s_2)$ for formulation Φ, and thus the approximation is very close for large t and $q = 1$. For formulation Ψ, we can write (18) in the following manner:

$$\frac{\pi_\Psi(a)}{\eta(a_u)} = \frac{\pi_\Phi(a)}{(1-q)\eta(a_u)} - \frac{qc_r}{1-q}. \tag{21}$$

For q very close to one, $\pi_\Phi(a)/\eta(a_u) = c_r$ and thus $\pi_\Psi(a)/\eta(a_u) = c_r((1/(1-q)) - (q/(1-q))) = c_r$.

From (18), $c_r - (\pi_\Psi(a)/\eta(a_u)) = (\pi_\Phi(a) - \pi_\Psi(a))/q\eta(a_u)$. For lower values of q—for instance, $q = 0.01$—we have from Table I the ∞-norm of $\pi_\Phi - \pi_\Psi$ to be in the range of $3.71E{-}4$, which when divided by q will be of the range .037. As $0 \leq \eta(a_u) \leq 1$, the difference in c_r and $\pi_\Psi(a)/\eta(a_u)$ can be substantial for lower qs.

A better approximation of the transition probabilities of the reduced network ϕ_Ψ (context-reduced network from Ψ) can be derived by using $P(w_t = 2^n j + s_2) = \pi_\Psi(2^n j + s_2)$ in (19). The resulting state transition probabilities of the context-reduced network is given by (22), shown at the bottom of the page. Let the steady-state probability distribution of the context-reduced PBN ϕ_Ψ be denoted by the vector ϑ. The next theorem shows that if we use (22) for calculating the transition-probabilities of the context-reduced PBN; then the steady-state probability distribution vector ϑ of the context-reduced PBN will be the same as the collapsed steady-state probability distribution vector η.

Theorem III.1: With (22) denoting the transition probabilities of the reduced PBN, the steady-state probability distribution vector ϑ of the reduced PBN is given by

$$\vartheta(j) = \sum_{i=1}^{k}\pi_\Psi\left(2^n(i-1)+j\right) = \eta_\Psi(j) \tag{23}$$

for $j = 0, 1, \ldots, 2^n - 1$.

Proof: η_Ψ is the steady-state probability distribution of the reduced PBN if it satisfies the equation $\sum_{i=0}^{2^n-1}\eta_\Psi(i)P_{\phi_\Psi}(i,j) = \eta_\Psi(j)$ for each $j \in 0, 1, \ldots, 2^n - 1$.

Since π is the steady-state probability distribution for the $2^n k \times 2^n k$ PBN Ψ, it satisfies $\pi_\Psi(l_2) = \sum_{l_1=0}^{2^n k-1}\pi_\Psi(l_1)P_\Psi(l_1, l_2)$. Thus

$$\begin{aligned}
\eta_\Psi(j) &= \sum_{i=1}^{k}\pi_\Psi\left(2^n(i-1)+j\right) \\
&= \sum_{i=1}^{k}\sum_{l_1=0}^{2^n k-1}\pi_\Psi(l_1)P_\Psi\left(l_1, 2^n(i-1)+j\right) \\
&= \sum_{i=1}^{k}\sum_{j_2=0}^{2^n-1}\sum_{i_2=1}^{k}\pi_\Psi\left(2^n(i_2-1)+j_2\right)P_\Psi \\
&\quad \times \left(2^n(i_2-1)+j_2, 2^n(i-1)+j\right) \\
&= \sum_{j_2=0}^{2^n-1}\left(\sum_{i_3=1}^{k}\pi_\Psi\left(2^n(i_3-1)+j_2\right)\right)P_{\phi_\Psi}(j_2, j) \\
&= \sum_{j_2=0}^{2^n-1}\eta_\Psi(j_2)P_{\phi_\Psi}(j_2, j).
\end{aligned}$$

$$\begin{aligned}
P_\phi(s_1,s_2) &= P(z_{t+1} = s_2 | z_t = s_1) \\
&= \sum_{i=1}^{k} P(z_{t+1} = s_2, s_2 \text{ in network } i | z_t = s_1) \\
&= \frac{\sum_{i=1}^{k} P(z_{t+1} = s_2, s_2 \text{ in network } i, z_t = s_1)}{P(z_t = s_1)} \\
&= \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} P(z_{t+1} = s_2, z_t = s_1, s_2 \text{ in net } i, s_1 \text{ in net } j)}{\sum_{j'=1}^{k} P(z_t = s_1, s_1 \text{ in network } j')} \\
&= \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} P\left(w_{t+1} = 2^n(i-1)+s_2 | w_t = 2^n(j-1)+s_1\right) P\left(w_t = 2^n(j-1)+s_1\right)}{\sum_{j'=1}^{k} P\left(w_t = 2^n(j'-1)+s_1\right)}
\end{aligned} \tag{19}$$

$$P_{\phi_\Psi}(s_1,s_2) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} P\left(w_{t+1} = 2^n(i-1)+s_2 | w_t = 2^n(j-1)+s_1\right)\pi_\Psi\left(2^n(j-1)+s_1\right)}{\sum_{j'=1}^{k}\pi_\Psi\left(2^n(j'-1)+s_1\right)} \tag{22}$$

This shows that η_Ψ satisfies the steady-state probability equations of the reduced network and is equal to ϑ due to the uniqueness of the steady-state probability distribution for ergodic Markov chains.[6]

Corollary III.2: If the context-reduced transition probabilities for formulation Φ is given by (24), shown at the bottom of the page, for $s_1 = 0, 1, \ldots, 2^n - 1$ and $s_2 = 0, 1, \ldots, 2^n - 1$, then the steady-state probability distribution of ϕ_Φ denoted by ϑ_Φ is given by

$$\vartheta_\Phi(j) = \sum_{i=1}^{k} \pi_\Phi\left(2^n(i-1)+j\right) = \eta_\Phi(j) \tag{25}$$

for $j = 0, 1, \ldots, 2^n - 1$.

As shown in (23) and (25), the new reduction is capable of matching the steady-state probability distribution of the context-reduced PBN to the collapsed steady-state probability distribution of the original context-sensitive PBN, which was not feasible in the earlier approximation [26]. In the earlier approximation, the context-reduced PBN was independent of q and was equivalent to $q = 1$ in the current case. So the current formulation will better capture the dependence on q.

IV. Approximation Method for Calculating the Steady-State Probabilities of the Context-Sensitive PBN

To explain the importance of the calculation of the steady-state distribution for both context and gene states, let us consider the dynamical modeling of cells in a tumor being treated by a targeted molecular therapeutic drug. Targeted cancer therapy refers to a new generation of cancer drugs designed to interfere with a specific molecular target that is believed to have a critical role in tumor growth or progression [38], [39]. However, resistance to the drug can evolve due to the activation of alternative pathways [40]. The resistance to a specific targeted drug can disappear with stopping of that drug therapy or use of another targeted drug. Modeling of such a scenario with a context-sensitive probabilistic Boolean network is desirable, as the drug-sensitive and drug-resistant scenarios can be considered as two separate contexts with different network connections in each scenario, and they can switch between themselves with small probabilities. The two scenarios can be further divided into more contexts referring to the different signaling pathway activations for each individual tumor. The GAP states here will consist of the candidate gene or protein expression values important for the tumor. We will be interested in knowing the steady-state probability of each GAP state in the resistant and sensitive contexts, with an immediate application being a high steady-state probability of a GAP state in one of the contexts and relatively lower steady-state probability in the other context, can signify the suitability of the GAP state as a molecular marker for drug resistance/sensitivity.

Previous approximation methods for calculation of steady-state probabilities of a PBN have considered generation of the steady-state probabilities of the major attractors of the constituent Boolean networks [25] or removal of Boolean networks with small selection probabilities [32]. In this paper, we will approach the approximation based on the steady-state probability distribution of the constituent Boolean networks and the transitions between them. Instead of calculating the steady-state probability distribution of the context-sensitive PBN directly with a $2^n k \times 2^n k$ transition probability matrix, we will calculate the steady-state probability distributions of the k constituent networks with $2^n \times 2^n$ transition probability matrices and utilize the frequency of transitions between the constituent networks to arrive at the approximation. The proposed method provides good approximation for small values of network transition probability q.

Here we will consider only π_Ψ, knowing that π_Φ can be easily calculated from π_Ψ using (18). The transition probabilities for BN l with perturbation are given by the $2^n \times 2^n$ matrix P_l. The individual terms of P_l are given by $P_l(i,j) = (1-p)^n f_{l,i,j} + (1-p)^{n-h} p^h s(h)$ for $i, j \in 0, 1, \ldots, 2^n - 1$, where p is the perturbation probability, h is the Hamming distance between i and j, and the functions f and s are given by (2) and (4), respectively. Let the steady-state probability distribution for network l be denoted by a 2^n-dimensional vector ρ_l. To calculate the steady-state probability $\pi_\Psi(2^n(i-1)+j)$, we will consider the four mutually exclusive events $\Psi 1$ to $\Psi 4$ described earlier. The probability corresponding to events $\Psi 1$ and $\Psi 3$ can be approximated by $c_i(1-q+qc_i)\rho_i(j)$. The approximation is due to the fact that the probability of being in network i is c_i; the probability of staying in network i at the next time step is $1-q+qc_i$; and the probability of GAP state j in network i can be approximated by $\rho_i(j)$.

To approximate the probability corresponding to events $\Psi 2$ and $\Psi 4$, let us consider the event that a network change from i_2 to i has occurred. Then according to formulation Ψ, the probability of occurrence of a transition from state $2^n(i_2-1)+j_3$ to the state $2^n(i-1)+j$ where $i_2 \neq i$ will be given by the following equation:

$$\begin{aligned} &P\left(w(t) = 2^n(i_2-1)+j_3 \text{ and } w(t+1) = 2^n(i-1)+j\right) \\ &\quad = P\left(w(t+1) = 2^n(i-1)+j \middle| w(t) = 2^n(i_2-1)+j_3\right) \\ &\qquad \times P\left(w(t) = 2^n(i_2-1)+j_3\right) \\ &\quad \approx P(\text{transition from net } i_2 \text{ to net } i) \\ &\qquad \times P(\text{GAP transition from } j_3 \text{ to } j \text{ in net } i) \\ &P\left(w(t) = 2^n(i_2-1)+j_3\right) \\ &\quad \approx \left(qc_i\right)\left(P_i(j_3,j)\right)\left(c_{i_2}\rho_{i_2}(j_3)\right). \end{aligned} \tag{26}$$

[6]The nonzero perturbation probability p ensures the ergodicity of a context-sensitive PBN.

$$P_{\phi_\Phi}(s_1,s_2) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} P\left(w_{t+1} = 2^n(i-1)+s_2 | w_t = 2^n(j-1)+s_1\right)\pi_\Phi\left(2^n(j-1)+s_1\right)}{\sum_{j'=1}^{k}\pi_\Phi\left(2^n(j'-1)+s_1\right)} \tag{24}$$

Here, we have used the approximation that the probability of state $2^n(i_2-1)+j_3$ in the steady state is $c_{i_2}\rho_{i_2}(j_3)$.

As the network i_2 in the above equation can be any network from network 1 to network k except network i, and the state j_3 for each i_2 can range from zero to 2^n-1, the steady-state probability for $2^n(i-1)+j$ can be approximated by the following equation:

$$\pi_{\Psi_{ap}}(2^n(i-1)+j) = c_i(1-q+qc_i)\rho_i(j) + \sum_{i_2=1, i_2\neq i}^{k}\sum_{j_3=0}^{2^n-1} qc_{i_2}c_iP_i(j_3,j)\rho_{i_2}(j_3). \quad (27)$$

The utility of the above approximation lies in the fact that we can compute the steady-state probabilities of a context-sensitive PBN with a reduced computational complexity. The steady-state probabilities of a $2^nk \times 2^nk$ matrix can be computed by a matrix inversion with a complexity of $O(2^{3n}k^3)$ using Gaussian elimination.[7] In the new approach, we can calculate the steady-state probabilities for $k2^n \times 2^n$ networks with a complexity of $O(k2^{3n})$. The rest of the summation calculations are of lower order $(O(2^nk))$. For large k, the reduction in complexity from $O(2^{3n}k^3)$ to $O(k2^{3n})$ will be significant. We next show analytically and through simulations that the approximation is quite close to the original solution for smaller values of q.

A. Bound on the Approximation

The steady-state probability distribution π_Ψ of P_Ψ is the solution to the following linear equation:

$$x(P_\Psi - I_{2^nk}) = \mathbf{0} \quad (28)$$

where I_{2^nk} is the identity matrix of size $2^nk \times 2^nk$ and the solution vector x is equal to π_Ψ. The residual for an approximate solution π_a of (28) is given by $\beta = \pi_a(P_\Psi - I_{2^nk})$, where the norm of β provides a measure of closeness of the approximate solution as compared to the true solution [42].

Let us first study the closeness of the approximate solution $\pi_a = (c_1\rho_1, c_2\rho_2, \ldots, c_k\rho_k)$, where ρ_i is the 2^n dimensional steady-state probability distribution vector for P_i. The following theorem provides a bound on the infinity norm of the residual for the approximate solution π_a for (28).

Theorem IV.1: $|\beta|_\infty \leq q\max(c_1, c_2, \ldots, c_k)$, where β is the residual corresponding to the approximate solution $\pi_a = (c_1\rho_1, c_2\rho_2, \ldots, c_k\rho_k)$ for (28).

Proof:

$$\begin{aligned} P_\Psi - I_{2^nk} &= \begin{bmatrix} (1-q+qc_1)P_1 & qc_2P_2 & \cdots & qc_kP_k \\ \cdots & \cdots & \cdots & \cdots \\ qc_1P_1 & qc_2P_2 & \cdots & (1-q+qc_k)P_k \end{bmatrix} - I_{2^nk} \\ &= (1-q)\begin{bmatrix} P_1 - I_{2^n} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & P_k - I_{2^n} \end{bmatrix} \\ &\quad + q\begin{bmatrix} c_1P_1 - I_{2^n} & c_2P_2 & \cdots & c_kP_k \\ \cdots & \cdots & \cdots & \cdots \\ c_1P_1 & c_2P_2 & \cdots & c_kP_k - I_{2^n} \end{bmatrix}. \end{aligned} \quad (29)$$

[7] An improved complexity of $O(N^{2.376})$ can be achieved by the Coppersmith and Winograd method [41].

As $\rho_i(P_i - I_{2^n}) = \rho_i - \rho_i = \mathbf{0}$ for $i \in [1, 2, \ldots, k]$

$$\begin{aligned} \beta &= \pi_a(P_\Psi - I_{2^nk}) \\ &= (1-q)\mathbf{0} + q|c_1\chi P_1 - c_1\rho_1,\ c_2\chi P_2 - c_2\rho_2,\ \cdots, \\ &\qquad c_k\chi P_k - c_k\rho_k| \end{aligned}$$

where $\chi = \sum_{i=1}^{k} c_i\rho_i$.

As the multiplication of a state probability vector with a transition probability matrix produces another state probability vector, ρ_iP_js are state probability vectors for $i \in [1, \ldots, k]$ and $j \in [1, \ldots, k]$. Since $\sum_{i=1}^{k} c_i = 1$, $|\chi P_i|_\infty \leq 1$ for $i = 1, 2, \ldots, k$. Furthermore, as individual entries of χP_i and ρ_is are positive and ≤ 1, $c_i|\chi P_i - \rho_i|_\infty \leq c_i$. Thus, $q|c_1\chi P_1 - c_1\rho_1, c_2\chi P_2 - c_2\rho_2, \ldots, c_k\chi P_k - c_k\rho_k|_\infty \leq q\max(c_1, c_2, \ldots, c_k)$.

Corollary IV.2: $|\beta|_1 = q|c_1\chi P_1 - c_1\rho_1, c_2\chi P_2 - c_2\rho_2, \ldots, c_k\chi P_k - c_k\rho_k|_1 \leq 2q$.

Let us now consider the approximation provided by (27)

$$\begin{aligned} &\pi_{\Psi_{ap}}(2^n(i-1)+j) \\ &\quad = c_i(1-q+qc_i)\rho_i(j) + \sum_{i_2=1, i_2\neq i}^{k}\sum_{j_3=0}^{2^n-1} qc_{i_2}c_iP_i(j_3,j)\rho_{i_2}(j_3) \\ &\quad = (1-q)c_i\rho_i(j) + qc_i\sum_{i_2=1}^{k}\sum_{j_3=0}^{2^n-1} c_{i_2}P_i(j_3,j)\rho_{i_2}(j_3) \\ &\qquad \left(\text{As } \sum_{j_3=0}^{2^n-1} P_i(j_3,j)\rho_i(j_3) = \rho_i(j)\right) \\ &\quad = \sum_{j_4=0}^{2^nk-1} \pi_a(j_4)P_\Psi(j_4, 2^n(i-1)+j) \\ &\qquad (\text{due to (11) and as } \pi_a(2^n(i-1)+j) = c_i\rho_i(j)). \end{aligned}$$

Thus the steady-state probability approximation $\pi_{\Psi_{ap}}$ provided by (27) is equal to π_aP_Ψ. Consequently, the residual for the approximation is $\beta_2 = \pi_aP_\Psi(P_\Psi - I_{2^nk}) = (\beta+\pi_a)(P_\Psi - I_{2^nk}) = rP_\Psi - \beta + \pi_aP_\Psi - \pi_a = \beta P_\Psi$

$$\begin{aligned} \beta P_\Psi &= \beta\left((1-q)\begin{bmatrix} P_1 & 0 & \cdots \\ \cdots & \cdots & \cdots \\ 0 & \cdots & P_k \end{bmatrix} + q\begin{bmatrix} c_1P_1 & c_2P_2 & \cdots \\ \cdots & \cdots & \cdots \\ c_1P_1 & \cdots & c_kP_k \end{bmatrix}\right) \\ &= q(1-q)|c_1\chi P_1 - c_1\rho_1, \ldots, c_k\chi P_k - c_k\rho_k| \\ &\quad \times \begin{bmatrix} P_1 & 0 & \cdots \\ \cdots & \cdots & \cdots \\ 0 & \cdots & P_k \end{bmatrix} \\ &\quad + q^2|c_1\chi P_1 - c_1\rho_1, \ldots, c_k\chi P_k - c_k\rho_k| \\ &\quad \times \begin{bmatrix} c_1P_1 & c_2P_2 & \cdots \\ \cdots & \cdots & \cdots \\ c_1P_1 & \cdots & c_kP_k \end{bmatrix} \\ &= q(1-q)\left|c_1\left(\chi P_1^2 - \rho_1P_1\right), \ldots, c_k\left(\chi P_k^2 - \rho_kP_k\right)\right| \\ &\quad + q^2|c_1(\chi\Upsilon P_1 - \chi P_1), \ldots, c_k(\chi\Upsilon P_k - c_k\chi P_k)| \\ &\quad \left(\text{where } \Upsilon = \sum_{i=1}^{k} c_iP_i\right). \end{aligned} \quad (30)$$

TABLE II
APPROXIMATION ERRORS FOR DIFFERENT SETS OF n, p, q, k, WHERE $E = \pi_\Psi - \pi_{\Psi_{ap}}$

n	k	p	q	$\|\|E\|\|_1$	$\|\|E\|\|_2$	$\|\|E\|\|_\infty$	$\|\|\beta P_\Psi\|\|_1$	$\|\|\beta P_\Psi\|\|_\infty$	t_{or}	t_{ap}
4	15	.01	.01	3.32E-2	6.90E-3	4.14E-3	6.39E-3	5.73E-4	0.01	0.0049
4	15	.05	.05	3.66E-2	5.63E-3	2.89E-3	1.71E-3	1.17E-3	0.01	0.0045
4	15	.01	.1	1.51E-1	2.83E-2	1.64E-2	5.65E-2	4.86E-3	0.01	0.0045
4	15	.25	.05	1.37E-3	1.61E-4	6.97E-5	1.17E-3	5.5E-5	0.01	0.0048
5	15	.01	.01	3.53E-2	5.58E-3	3.13E-2	7.64E-3	5.12E-4	0.06	0.013
5	15	.01	.1	1.92E-1	2.68E-2	1.50E-2	6.72E-2	4.11E-3	0.06	0.013
5	15	.01	.5	3.73E-1	4.65E-2	2.46E-2	2.39E-1	1.37E-2	0.061	0.013
5	15	.25	.01	1.88E-4	1.47E-5	4.41E-6	1.45E-4	3.33E-6	0.06	0.013
5	15	.5	.01	4.50E-7	3.56E-8	1.12E-8	4.72E-7	1.18E-8	0.06	0.013
6	15	.01	.01	3.79E-2	4.66E-3	2.71E-3	8.01E-3	3.99E-4	0.42	0.044
7	15	.01	.01	4.17E-2	3.97E-3	2.25E-3	8.11E-3	2.9E-4	3.01	0.164
8	15	.01	.01	4.25E-2	3.18E-3	1.64E-3	7.86E-3	1.78E-4	22.35	0.677
8	15	.05	.05	2.05E-2	8.34E-4	2.65E-4	1.06E-2	9.6E-5	22.33	0.678
8	5	.05	.05	1.70E-2	1.05E-3	3.26E-4	9.05E-3	1.25E-4	0.956	0.109

As individual entries of χ are positive and the sum of the entries of χ add up to one, χ is a valid state probability distribution vector. Since multiplication of a state probability distribution vector with a probability transition matrix results in a state probability distribution vector, χP_i^2 is a state probability distribution vector. Since the entries of Υ are nonnegative and rows sum to one, $\chi \Upsilon P_i$ is a state probability distribution vector. Thus, $|\chi P_i^2 - \rho_i P_i|_\infty \leq 1$ and $|\chi \Upsilon P_i - \chi P_i|_\infty \leq 1$ for $i \in 1, 2, \ldots, k$. Based on these inequalities, we can get a bound on $|\beta P_\Psi|_\infty$ as follows:

$$\begin{aligned}|\beta P_\Psi|_\infty &\leq q(1-q)\left|c_1\left(\chi P_1^2 - \rho_1 P_1\right), \cdots, c_k\left(\chi P_k^2 - \rho_k P_k\right)\right|_\infty \\ &\quad + q^2 \left|c_1(\chi \Upsilon P_1 - \chi P_1), \ldots, c_k(\chi \Upsilon P_k - c_k \chi P_k)\right|_\infty \\ &\leq q(1-q)\max(c_1, c_2, \cdots, c_k) + q^2 \max(c_1, c_2, \cdots, c_k) \\ &\leq q \max(c_1, c_2, \ldots, c_k). \end{aligned} \tag{31}$$

Similarly

$$\begin{aligned}|\beta P_\Psi|_1 &\leq q(1-q)\left|c_1\left(\chi P_1^2 - \rho_1 P_1\right), \ldots, c_k\left(\chi P_k^2 - \rho_k P_k\right)\right|_1 \\ &\quad + q^2 \left|c_1(\chi \Upsilon P_1 - \chi P_1), \ldots, c_k(\chi \Upsilon P_k - c_k \chi P_k)\right|_1 \\ &\leq q(1-q) 2(c_1 + c_2 + \cdots + c_k) \\ &\quad + q^2 2(c_1 + c_2 + \cdots + c_k) \\ &\leq 2q \end{aligned} \tag{32}$$

B. Simulation Results

In this section, we present the results of simulations conducted to measure the accuracy of the approximation algorithm given by (27). Let E denote the difference between the true steady-state probability distribution and the approximate distribution. The error measures used are the $\|.\|_1$, $\|.\|_2$, and $\|.\|_\infty$ norms of E and the $\|.\|_1$ and $\|.\|_\infty$ norms of βP_Ψ (residual for the approximate solution). We considered different values of n, k, p, and q and constructed 50 context-sensitive PBNs for each set of n, k, p, and q values. The average error of the 50 context-sensitive PBNs for specific sets of n, k, and q is shown in Table II. The detailed table containing a large number of variations of n, k, p, and q is provided in the companion Web site.

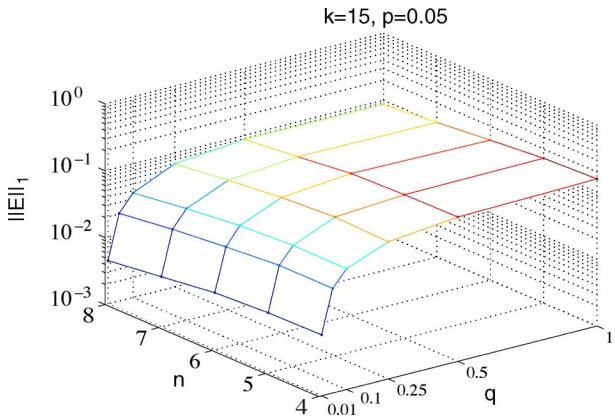

Fig. 4. Average $\|E\|_1$ for $k = 15$, $p = 0.05$, and different values of n and q.

The simulations were conducted using Matlab on a Dell Optiplex GX745 with a 2.4 GHz Intel Core2 CPU and 2 GB of RAM. In the table, t_{or} denotes the time in seconds taken to compute the true steady-state probability distribution of the context-sensitive PBN and t_{ap} denotes the time in seconds taken by the approximation algorithm to compute the steady-state probability distribution of the context-sensitive PBN. We observe that the approximation is close to the true steady-state probability distribution for low values of q. Whenever $q \leq p$, the 1-norm of E is less than 0.045. The 1-norm of E for different values of n and q and fixed $k = 15$ and $p = 0.05$ is plotted in Fig. 4. The figure shows that the 1-norm of the error for the approximation algorithm increase when q increases. An analogous relationship

holds for the 2-norm and the ∞-norm. The simulations show that the 1-norm of the residual $|\beta P_\Psi|_1$ is always $\leq 2q$, as is expected from (32). When we study the relationship of the error with changes in p, the error decreases with increase in the value of p. The more random the individual networks become, the approximation algorithm has better performance. For instance, as shown in Table II, for $n = 5$, $k = 15$, $p = .5$, and $q = .01$, the 1-norm of the error is $4.5 * 10^{-7}$, which is very low. In terms of time complexity, we observe that for large k, the time taken by the approximation algorithm is much less than the time taken to calculate the actual steady-state probability distribution. The entry corresponding to $n = 8$, $p = .05$, $q = .05$, and $k = 15$ in Table II shows the advantage of the approximation algorithm in terms of time complexity. The true steady-state probability distribution is computed in 22.33 s, but the approximation can be generated in only 0.678 s.

V. Conclusion

In this paper, we considered three issues related to context-sensitive PBNs. First, we derived the relationship between the steady-state probability distributions of the two context-sensitive formulations with different ordering of the same set of events. This will facilitate better understanding of the context-sensitive probabilistic Boolean network formulation. We next derived a mapping for removal of context information from a context-sensitive PBN that maintains the collapsed steady-state probability distribution. For our mapping, we analytically proved the equality of the steady-state probability distribution of the reduced PBN and the collapsed steady-state probability distribution of the original PBN. As the steady state is a crucial characteristic of the regulatory network that denotes the phenotype, this reduction algorithm will be valuable for preserving the steady-state probability distribution after reduction. Lastly, we proposed an approximation method for the calculation of the steady-state probabilities for a context-sensitive PBN, which is significantly faster than the calculation of the true steady-state probability distribution. Future research issues in this area include combining the context-reduction with the downstream control design with the objective of minimum alteration on the control outcome.

References

[1] H. De Jong, "Modeling and simulation of genetic regulatory systems: A literature review," *J. Comput. Biol.*, vol. 9, pp. 67–103, 2001.

[2] I. Shmulevich and E. R. Dougherty, *Genomic Signal Processing*. Princeton, NJ: Princeton Univ. Press, 2007.

[3] J. M. Bower and H. Bolouri, *Computational Modeling of Genetic and Biochemical Networks*, 1st ed. Boston, MA: MIT Press, 2001.

[4] R. Sharan and T. Ideker, "Modeling cellular machinery through biological network comparison," *Nat. Biotechnol.*, vol. 24, pp. 427–433, 2006.

[5] E. D. Conrad and J. J. Tyson, "Modeling molecular interaction networks with nonlinear ordinary differential equations," in *System Modeling in Cell Biology From Concepts to Nuts and Bolts*. Cambridge, MA: MIT Press, 2006, pp. 97–123.

[6] D. S. Douglas and S. Jones, *Differential Equations and Mathematical Biology*, 1st ed. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[7] L. Qian, H. Wang, and E. Dougherty, "Inference of noisy nonlinear differential equation models for gene regulatory networks using genetic programming and Kalman filtering," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3327–3339, 2008.

[8] M. de Hoon, S. Imoto, and S. Miyano, *Inferring Gene Regulatory Networks From Time-Ordered Gene Expression Data Using Differential Equations*. Berlin, Germany: Springer, 2002.

[9] J. Goutsias and S. Kim, "Stochastic transcriptional regulatory systems with time delays: A mean-field approximation," *J. Comput. Biol.*, vol. 13, pp. 1049–1076, 2006.

[10] J. Goutsias and S. Kim, "A nonlinear discrete dynamical model for transcriptional regulation: Construction and properties," *Biophys. J.*, vol. 86, pp. 1922–1945, 2004.

[11] A. Margolin, I. Nemenman, K. Basso, U. Klein, C. Wiggins, G. Stolovitzky, R. D. Favera, and A. Califano, "Aracne: An algorithm for reconstruction of genetic networks in a mammalian cellular context," *BMC Bioinf.*, vol. 7, 2006.

[12] L. Glass and S. A. Kauffman, "The logical analysis of continuous, nonlinear biochemical control networks," *J. Theor. Biol.*, vol. 39, pp. 103–129, 1973.

[13] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," *Theor. Biol.*, vol. 22, pp. 437–467, 1969.

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

[15] R. Somogyi and C. Sniegoski, "Modeling the complexity of gene networks: Understanding multigenic and pleiotropic regulation," *Complexity*, vol. 1, pp. 45–63, 1996.

[16] Z. Szallasi and S. Liang, "Modeling the normal and neoplastic cell cycle with realistic boolean genetic networks: Their application for understanding carcinogenesis and assessing therapeutic strategies," *Proc. Pac. Symp. Biocomput.*, vol. 3, pp. 66–76, 1998.

[17] R. Thomas, D. Thieffry, and M. Kaufman, "Dynamical behavior of genetic regulatory networks i. biological role of feedback loops and practical use of the concept of the loop-characteristic state," *Bull. Math. Biol.*, vol. 57, pp. 247–276, 1995.

[18] A. Wuensche, "Genomic regulation modeled as a network with basins of attraction," *Proc. Pac. Symp. Biocomput.*, vol. 3, pp. 89–102, 1998.

[19] S. Huang, "Gene expression profiling, genetic networks, and cellular states: An integrating concept for tumorigenesis and drug discovery," *Molec. Med.*, vol. 77, pp. 469–480, 1999.

[20] A. Faure, A. Naldi, C. Chaouiya, and D. Theiffry, "Dynamical analysis of a generic boolean model for the control of the mammalian cell cycle," *Bioinformatics*, vol. 22, pp. 124–131, 2006.

[21] J. C. Sible and J. J. Tyson, "Mathematical modeling as a tool for investigating cell cycle control networks," *Methods*, vol. 41, pp. 238–247, 2007.

[22] J. J. Tyson, A. Csikasz-Nagy, and B. Novak, "The dynamics of cell cycle regulation," *Bioessays*, vol. 24, pp. 1095–1109, 2002.

[23] 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, pp. 261–274, 2002.

[24] I. Shmulevich, E. R. Dougherty, and W. Zhang, "From boolean to probabilistic boolean networks as models of genetic regulatory networks," *Proc. IEEE*, vol. 90, pp. 1778–1792, 2002.

[25] M. Brun, E. R. Dougherty, and I. Shmulevich, "Steady-state probabilities for attractors in probabilistic Boolean networks," *Signal Process.*, vol. 85, pp. 1993–2013, 2005.

[26] R. Pal, A. Datta, M. L. Bittner, and E. R. Dougherty, "Intervention in context-sensitive probabilistic boolean networks," *Bioinformatics*, vol. 21, pp. 1211–1218, 2005.

[27] B. Faryabi, G. Vahedi, J.-F. Chamberland, A. Datta, and E. R. Dougherty, "Intervention in context-sensitive probabilistic boolean networks revisited," *EURASIP J. Bioinf. Syst. Biol.*, vol. 2009, p. ID360864, 2009.

[28] R. Pal, A. Datta, and E. R. Dougherty, "Robust intervention in probabilistic Boolean networks," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1280–1294, Mar. 2008.

[29] I. Ivanov and E. R. Dougherty, "Reduction mappings between probabilistic boolean networks," *EURASIP J. Appl. Signal Process.*, vol. 4, pp. 125–131, 2004.

[30] I. Ivanov, R. Pal, and E. R. Dougherty, "Dynamics-preserving size-reduction mappings for probabilistic Boolean networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2310–2322, May 2007.

[31] E. R. Dougherty and I. Shmulevich, "Mappings between probabilistic boolean networks," *Signal Process.*, vol. 83, pp. 799–809, 2003.

[32] 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, pp. 1511–1518, 2007.
[33] R. Layek, A. Datta, R. Pal, and E. R. Dougherty, "Adaptive intervention in probabilistic boolean networks," in *Proc. Amer. Contr. Conf.*, accepted for publication.
[34] A. Datta, A. Choudhary, M. L. Bittner, and E. R. Dougherty, "External control in Markovian genetic regulatory networks," *Mach. Learn.*, vol. 52, pp. 169–191, 2003.
[35] A. Datta, A. Choudhary, M. L. Bittner, and E. R. Dougherty, "External control in Markovian genetic regulatory networks: The imperfect information case," *Bioinformatics*, vol. 20, pp. 924–930, 2004.
[36] R. Pal, A. Datta, and E. R. Dougherty, "Optimal infinite horizon control for probabilistic Boolean networks," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2375–2387, 2006.
[37] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Nashua, NH: Athena Scientific, 2001.
[38] C. Sawyers, "Targeted cancer therapy," *Nature*, vol. 432, pp. 294–297, 2004.
[39] M. R. Green, "Targeting targeted therapy," *New Eng. J. Med.*, vol. 350, no. 21, pp. 2191–2193, 2004.
[40] N. V. Zhukov and S. A. Tjulandin, "Targeted therapy in the treatment of solid tumors: Practice contradicts theory," *Biochemistry (Mosc).*, vol. 73, p. 605-188, 2008.
[41] D. Coppersmith and S. Winograd, "Matrix multiplication via arithmetic progressions," *J. Symbol. Comput.*, vol. 9, pp. 251–280, 1990.
[42] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

Ranadip Pal (M'07) received the B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, in 2002 and the M.S. and Ph.D. degrees in electrical engineering from Texas A&M University, College Station, in 2004 and 2007, respectively.

In August 2007, he joined the Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, as an Assistant Professor. His research areas are computational biology, genomic signal processing, and control of genetic regulatory networks.

Dr. Pal is the recipient of the 2010 NSF CAREER Award. While with Texas A&M, Dr. Pal received the Ebensberger/Fouraker Fellowship, a Distinguished Graduate Student Masters Research Award, and a National Instruments Fellowship. He was an Indian National Math Olympiad Awardee and was ranked first in the Regional Math Olympiad, West Bengal, India.