

State Estimation for Stochastic Time-Varying Boolean Networks

Hongwei Chen, Zidong Wang, Jinling Liang and Maozhen Li

Abstract—In this paper, a general theoretical framework is developed for the state estimation problem of stochastic time-varying Boolean networks (STVBNs). The STVBN consists of a system model describing the evolution of the Boolean states and a model relating the noisy measurements to the Boolean states. Both the process noise and the measurement noise are characterized by sequences of mutually independent Bernoulli distributed stochastic variables taking values of 1 or 0, which imply that the state/measurement variables may be flipped with certain probabilities. First, an algebraic representation of the STVBNs is derived based on the semi-tensor product. Then, based on the Bayes' theorem, a recursive matrix-based algorithm is obtained to calculate the one-step prediction and estimation of the forward/backward state probability distribution vectors. Owing to the Boolean nature of the state variables, the Boolean Bayesian filter is designed that can be utilized to provide the minimum mean-square error state estimate for the STVBNs. The fixed-interval smoothing filter is also obtained by resorting to the forward-backward technique. Finally, a simulation experiment is carried out for the context estimation problem of the *p53-MDM2* negative-feedback gene regulatory network.

Index Terms—Stochastic Boolean networks, state estimation, forward-backward smoothing, mean-square error, semi-tensor product.

I. INTRODUCTION

The concept of Boolean network (BN) has been proposed in the pioneering works by Kauffman [18] and Thomas [31]. BNs have emerged as prominent qualitative models of gene regulatory networks (GRNs) that, on one hand, require no kinetic information (or detailed mechanisms) and, on the other hand, unveil interesting dynamical phenomena of the network structure and the gene-gene interactions. In a BN model, the expression level of each gene is either “ON” or “OFF”, which is updated according to a preappointed Boolean function concerning the expression levels of other related genes. When such updates are performed for several time steps, trajectories of the BN are produced, which can then be utilized to compare with the experimental data for validation and refinement purposes. Notably, the qualitative nature of the states in the BN configuration makes it possible to simulate the dynamics of various biological systems [11] such as the cell cycle of several different organisms, the mammalian immune response to bacteria, the differentiation process of T-helper lymphocytes, as well as the GRN in *Drosophila melanogaster*.

Controlling the dynamics of GRNs is essential for developing new drugs and medical treatment techniques in the fields of biomedicine

This work was supported in part by the National Key R&D Program of China under Grant 2018YFB1003800, in part by the National Natural Science Foundation of China under Grant 61673110, Grant 61873148 and Grant 61933007, in part by the Shanghai Sailing Program under Grant 19YF1402400, in part by the Initial Research Funds for Young Teachers of Donghua University under Grant 104-07-0053019, and in part by the Alexander von Humboldt Foundation of Germany. (*Corresponding author: Jinling Liang.*)

H. Chen is with the College of Information Science and Technology, Donghua University, Shanghai 201620, China (e-mail: hongwei@dhu.edu.cn).

Z. Wang is with the Department of Computer Science, Brunel University London, Middlesex UB8 3PH, U.K. (e-mail: Zidong.Wang@brunel.ac.uk).

J. Liang is with the School of Mathematics, Southeast University, Nanjing 210096, China (e-mail: jinliang@seu.edu.cn).

M. Li is with the Department of Electronic and Computer Engineering, Brunel University London, Middlesex UB8 3PH, U.K. (e-mail: Maozhen.Li@brunel.ac.uk).

and bioengineering. An attractive abstraction, namely, Boolean control networks (BCNs), has been proposed in [1] for the sake of designing and analyzing the therapeutic intervention strategies. The ultimate objective of this line of research is to seek a discrete-time control sequence with minimum cost that will steer the network from a diseased state to a healthy one. Assisted by the semi-tensor product (STP) technique, a theoretic framework has been constructed for converting a BCN into its equivalent algebraic form that is quite similar to the standard discrete-time bilinear system [9]. This algebraic representation paves a convenient way for coping with the control-theoretic problems of BCNs, and many interesting results have been reported in the literature on various dynamics analysis problems such as stability and stabilization [3], [10], [23], [29], [39], [40], controllability and observability [26], [37], [38], optimal control [12], [19], [20], [35], system decomposition [28], [41], network synchronization [6], [24], [27]. In addition, the STP technique has been applied in tackling the problems of fault detection [13], asynchronous sequential machines [36], game theory [8] and so on.

Partially observed Boolean dynamical systems (POBDSs), originally introduced by Braga-Neto [4], have come to play a prominent role in many biological applications especially in genomic signal processing. Such models are nonlinear systems consisting of a stochastic Boolean model describing the state evolution and a noisy measurement model, thereby extending the classical Boolean models in two ways that account for i) uncertainty in the state evolution and ii) partial observation of the state variables through measurement noises. Many recent findings have shown that the POBDSs are capable of simulating the dynamical sequence of protein activation patterns of GRNs [2]. In practice, one of the main tasks is to figure out the true activity (state) of the genes based on the sample data. In most real-world applications, however, it is often the case that only partial information (or even no information) is observable about the states of practical GRNs through available measurement outputs. Hence, the state estimation issue for POBDSs has become a hot research topic with theoretical significance and practical importance, and many excellent results have been presented in the literature based on the Bayesian filtering approach [4], [16], [17].

Note that the state model and the measurement model in the proposed POBDSs are both governed by the Boolean function plus binary noise (i.e., modulo-2 addition), which implies that the introduced process and measurement noises are both “additive”. However, from the systems' point of view, the noise can appear in the form of any logical relationship related to the systems' state [7], [22], [25], [35], due mainly to the insufficient experimental evidence or incomplete understanding of a GRN. Recent results have further confirmed that the transitions between states in a GRN do occur in a random fashion [30], [32]–[34]. Theoretical models without taking the intrinsic stochastic noises into consideration might not be as serviceable as expected in practical applications. Therefore, stochastic noises should be recognized as a characteristic that has to be taken into account when modeling the dynamical behaviors of GRNs by BNs. In the remarkable paper [5], a new BN called stochastic time-invariant BN (STIBN) has been proposed, and the

inference of the discrete coefficient of determination for such kind of STIBN has been addressed. In recent years, many control-related problems of STIBN have been studied extensively, such as robust stabilization [22], robust control [25] and discounted optimal control [35]. On the other hand, it has now been well recognized that the dynamics of gene expression regulatory networks in practice is often time-varying [14], which leads to a general class of stochastic time-varying Boolean networks (STVBNs).

It is worth mentioning that all the aforementioned results on state estimation of POBDSs have been based on an implicit assumption that the process/measurement noises are additive, and therefore the developed filters cannot be directly utilized to deal with the general STVBNs. Specifically, the *prediction matrix* and the *update matrix* of the algorithms proposed in [4], [16] are both Boolean functions of the corresponding noise distribution, which are usually difficult to be calculated in case of the general STVBNs due to the lack of available tools to cope with the logical dynamic systems. As such, there is a practical need to investigate the state estimation problem for the general STVBNs. Such a problem is, to the best of the authors' knowledge, still open and remains challenging, and the purpose of this paper is therefore to shorten this gap by providing a rather rigorous framework.

The main objective of this paper is to design a recursive minimum-mean-square-error (MMSE) estimator for a class of BNs subjected to randomly occurring state/measurement flips. The main technical contributions of this paper can be highlighted as twofold. i) A recursive matrix-based algorithm is proposed to calculate the one-step prediction and estimation of the state probability distribution vector by resorting to the STP technique and Bayes' theorem, and Boolean Bayesian filter is also designed to estimate the true states of the STVBNs. ii) By applying the forward-backward approach, the fixed-interval smoothing filter is put forward to obtain the state estimate at each time instant based on a fixed interval of noisy measurements, where the estimation is usually better than that acquired by the Boolean Bayesian filter.

Notations: The notations used throughout this paper are fairly standard except where otherwise stated. \mathbb{N} and \mathbb{N}^+ denote the sets of nonnegative integers and positive integers, respectively. $\mathbb{R}^{m \times n}$ refers to the set of all $m \times n$ real matrices. The superscript "T" stands for matrix transposition. Δ_k means the set $\{\delta_k^i | i = 1, 2, \dots, k\}$, where δ_k^i is the i th column of the identity matrix I_k with degree k . A matrix $A \in \mathbb{R}^{m \times n}$ is called a logical matrix if $A = [\delta_m^{i_1} \delta_m^{i_2} \dots \delta_m^{i_n}]$ ($i_1, i_2, \dots, i_n \in \{1, 2, \dots, m\}$). $\mathcal{L}_{m \times n}$ is the set of all $m \times n$ logical matrices. Let $\text{Row}_i(A)$ (respectively, $\text{Col}_i(A)$) denote the i th row (respectively, column) of matrix A . The Bernoulli distribution with success probability p is simply denoted by $B(1, p)$. $\mathbb{E}\{\cdot\}$ is the mathematical expectation of a stochastic variable " \cdot ". $\|\cdot\|_1$ means the 1-norm of vectors.

II. MODEL FORMULATION

The STVBNs under consideration consist of a stochastic Boolean state model and a noisy measurement model. Specifically, the dynamic process of the state evolution is expressed as

$$X_i(k) = f_{k-1}^i(X(k-1), U(k-1), W(k-1)), \quad k \in \mathbb{N}^+ \quad (1a)$$

where $i = 1, 2, \dots, n$; $X(k) \triangleq (X_1(k), \dots, X_n(k))^T \in \mathcal{D}^n$ is the state variable with $\mathcal{D} \triangleq \{1, 0\}$; $U(k) \triangleq (U_1(k), \dots, U_m(k))^T \in \mathcal{D}^m$ stands for the control input which is assumed to be known; $W(k) \triangleq (W_1(k), \dots, W_{l_1}(k))^T \in \mathcal{D}^{l_1}$ represents the process noise obeying Bernoulli distribution with certain success probability, that is, $W_i(k) \sim B(1, p_k^i)$ for $i = 1, 2, \dots, l_1$; and $f_k^i(\cdot, \cdot, \cdot) : \mathcal{D}^{n+m+l_1} \rightarrow$

\mathcal{D} is the time-varying network function. Correspondingly, the measurement output can be described as follows:

$$Y_j(k) = h_k^j(X(k), V(k)), \quad j = 1, 2, \dots, q \quad (1b)$$

where, for $k \in \mathbb{N}^+$, $Y(k) \triangleq (Y_1(k), \dots, Y_q(k))^T \in \mathcal{D}^q$ ($q \leq n$) is the measurement information at discrete time k ; $V(k) \triangleq (V_1(k), \dots, V_{l_2}(k))^T \in \mathcal{D}^{l_2}$ represents the measurement noise with $V_j(k) \sim B(1, q_k^j)$ for $j = 1, 2, \dots, l_2$; and $h_k^j(\cdot, \cdot)$ is the time-varying Boolean function. In this paper, the noises $W(k)$ and $V(k)$ are assumed to be white and mutually independent. It is also supposed that the noise processes are independent with the initial state variable $X(0)$.

Remark 1: It is worth pointing out that, in Boolean model, the state space is not the usual continuous-valued set \mathbb{R}^n but a Boolean-valued one \mathcal{D}^n (i.e., $\{0, 1\}$). Although modern expression-based technologies (e.g., RNA-seq) would generate non-Boolean noises, the thresholding approaches can be utilized to binarize the noise expressions, which means Bernoulli-type process/measurement noises are possible to be realized/considered in this paper.

Remark 2: Note that the process and measurement noises introduced in [4] for POBDSs are both "additive", that is to say, both the state model and the measurement model are governed by the Boolean functions plus binary noises. While in the STVBN model (1), the noise can be in any form of the logical relationships with the system's states. From this point of view, system (1) under consideration is quite general, which includes the POBDS as a special case. In addition, the STVBN is capable of investigating some particular scenarios encountered in the transcriptomic analysis by allowing for noisy observations as well as incomplete measurements of the gene states (for instance, some of the genes in a pathway are not monitored).

Remark 3: Note that one can always assume that $p_k^i \leq \frac{1}{2}$ ($i = 1, 2, \dots, l_1$) and $q_k^j \leq \frac{1}{2}$ ($j = 1, 2, \dots, l_2$) without loss of generality. Specifically, if $p_k^i > \frac{1}{2}$, then one can employ an equivalent model in which $W_i(k)$ is substituted by $\overline{W_i(k)}$, where the overline " $\overline{\cdot}$ " represents the logical function "NOT", and the new noise parameter will have a success probability $1 - p_k^i \leq \frac{1}{2}$. On the other hand, it is worth pointing out that p_k^i and q_k^j represent the strength of the noises. If $p_k^i = q_k^j \equiv 0$ (or equivalently, 1) holds for all $k \in \mathbb{N}$, the network can be regarded as a noiseless one and the state transition can be viewed to be deterministic, while if $p_k^i = q_k^j = 0.5$, the uncertainty in the state evolution as well as output measurement is maximum, and for this particular case, the estimation discussed here does not make much sense. The estimation usually might have a better performance when p_k^i and q_k^j uniformly close to 0 with respect to k .

Given the noisy measurements $\{Y(1), \dots, Y(k)\}$ denoted by $Y_{1:k}$, the conditional mean-square error (MSE) of the estimator $\hat{X}(k)$ with respect to the unknown state variable $X(k)$ is defined as

$$\text{MSE}[\hat{X}(k)] \triangleq \mathbb{E} \left\{ \left\| X(k) - \hat{X}(k) \right\|_2^2 \middle| Y_{1:k}, U_{0:k-1} \right\} \quad (2)$$

where $U_{0:k-1} \triangleq \{U(0), \dots, U(k-1)\}$ and $\|\cdot\|_2$ means the Euclidean norm. Then, the *optimal state estimation problem* to be investigated in this paper is to find an estimator $\hat{X}(k) = \hat{g}(Y_{1:k}, U_{0:k-1})$ that minimizes the conditional MSE (2) at each time instant k for all possible realizations of $Y_{1:k}$ and $U_{0:k-1}$.

III. MAIN RESULTS

To facilitate the analysis of the optimal estimation problem, we first transform system (1) into its algebraic representation by utilizing the STP technique introduced in [9], where the Boolean variable $X_i(k) \in \mathcal{D}$ is identified with the vector $x_i(k) = \delta_2^{2-X_i(k)} \in \Delta_2$, and such a correspondence is extended naturally to the bijection between \mathcal{D}^n and

Δ_N with $N = 2^n$, i.e., the state vector $X(k)$ is regarded completely as the vector $x(k) = \times_{i=1}^n x_i(k) \in \Delta_N$. For each f_{k-1}^i (respectively, h_k^j), one can calculate its structure matrix \mathcal{F}_{k-1}^i (respectively, \mathcal{H}_k^j) [9], and system (1) can be converted into the following component-wise algebraic representation:

$$x_i(k) = \mathcal{F}_{k-1}^i w(k-1)u(k-1)x(k-1) \quad (3a)$$

$$y_j(k) = \mathcal{H}_k^j v(k)x(k) \quad (3b)$$

where $u(k) = \times_{i=1}^m u_i(k) \in \Delta_M$ with $M \triangleq 2^m$; $w(k) = \times_{i=1}^{L_1} w_i(k)$ and $v(k) = \times_{i=1}^{L_2} v_i(k)$ are, respectively, L_1 -valued and L_2 -valued random logical variables with $L_1 \triangleq 2^{l_1}$ and $L_2 \triangleq 2^{l_2}$. By further denoting $y(k) = \times_{j=1}^q y_j(k) \in \Delta_Q$ with $Q \triangleq 2^q$, using the properties of STP, the compact algebraic representation of system (1) can be obtained as follows:

$$x(k) = \mathcal{F}_{k-1} w(k-1)u(k-1)x(k-1) \quad (4a)$$

$$y(k) = \mathcal{H}_k v(k)x(k) \quad (4b)$$

where $\mathcal{F}_{k-1} = *_{i=1}^m \mathcal{F}_{k-1}^i \in \mathcal{L}_{N \times (ML_1N)}$ and $\mathcal{H}_k = *_{j=1}^q \mathcal{H}_k^j \in \mathcal{L}_{Q \times (L_2N)}$, in which “*” is the Khatri-Rao product.

A. Boolean Bayesian Filter

Suppose that the vector forms of the noisy measurements and the control sequence are $y_{1:k} \triangleq \{\delta_Q^{\lambda_1}, \dots, \delta_Q^{\lambda_k}\}$ and $u_{0:k-1} \triangleq \{\delta_M^{\mu_0}, \dots, \delta_M^{\mu_{k-1}}\}$, respectively. The forward state probability distribution vectors $\pi(k|j)$ are defined by

$$\pi(k|j) \triangleq \begin{bmatrix} \Pr\{x(k) = \delta_N^1 | y_{1:j}, u_{0:k-1}\} \\ \vdots \\ \Pr\{x(k) = \delta_N^N | y_{1:j}, u_{0:k-1}\} \end{bmatrix} \quad (5)$$

for $k = 1, 2, \dots$ and $j = k-1, k$, where $\Pr\{\cdot|\cdot\}$ refers to the conditional probability. In the notation that follows, we use $\pi(0|0)$ to denote the initial (or namely, prior) distribution of the state before any measurements are available. Note that $\pi(k|k-1)$ and $\pi(k|k)$ are the one-step prediction and the estimation of the forward state probability distribution vector at time instant k .

Based on the Bayes' theorem, the following result provides a recursive matrix-based algorithm to calculate the one-step prediction and estimation of the forward state probability distribution vector.

Lemma 1: Consider system (4) with noisy measurements $y_{1:k}$ and control sequence $u_{0:k-1}$. Let $\pi(0|0)$ be the initial distribution of the state. Then, we have

$$\pi(k|k-1) = (F_{k-1} \times \delta_M^{\mu_{k-1}}) \pi(k-1|k-1) \quad (6a)$$

$$\pi(k|k) = \frac{[\text{Row}_{\lambda_k}(H_k)]^\top \circ \pi(k|k-1)}{[\text{Row}_{\lambda_k}(H_k)] \pi(k|k-1)} \quad (6b)$$

where, for $k \in \mathbb{N}^+$, “ \circ ” represents the Hadamard (or element-wise) product and

$$F_{k-1} \triangleq \mathcal{F}_{k-1} \times_{i=1}^{L_1} [p_{k-1}^i \quad 1 - p_{k-1}^i]^\top \quad (7)$$

$$H_k \triangleq \mathcal{H}_k \times_{i=1}^{L_2} [q_k^i \quad 1 - q_k^i]^\top. \quad (8)$$

Proof: It follows from (5) that

$$\begin{aligned} \pi(k|k-1) &= \sum_{i=1}^N \delta_N^i \cdot \Pr\{x(k) = \delta_N^i | y_{1:k-1}, u_{0:k-1}\} \\ &= \mathbb{E}\{x(k) | y_{1:k-1}, u_{0:k-1}\} \end{aligned}$$

which, together with (4a), yields

$$\begin{aligned} \pi(k|k-1) &= \mathcal{F}_{k-1} \times \mathbb{E}\{w(k-1)\} \times u(k-1) \\ &\quad \times \mathbb{E}\{x(k-1) | y_{1:k-1}, u_{0:k-1}\} \\ &:= (F_{k-1} \times \delta_M^{\mu_{k-1}}) \pi(k-1|k-1) \end{aligned} \quad (9)$$

where $F_{k-1} = \mathcal{F}_{k-1} \mathbb{E}\{w(k-1)\} \in \mathbb{R}^{N \times MN}$ and

$$\begin{aligned} \mathbb{E}\{w(k-1)\} &= \times_{i=1}^{L_1} \mathbb{E}\{w_i(k-1)\} \\ &= \times_{i=1}^{L_1} [p_{k-1}^i \quad 1 - p_{k-1}^i]^\top. \end{aligned} \quad (10)$$

Now, we need to derive the measurement-update equation for the probability distribution vector $\pi(k|k)$. Based on the Bayes' theorem, the i th entry of $\pi(k|k)$, denoted by $\pi_i(k|k)$, can be obtained as

$$\begin{aligned} \pi_i(k|k) &= \frac{\Pr\{y(k)|x(k) = \delta_N^i\} \cdot \Pr\{x(k) = \delta_N^i | y_{1:k-1}, u_{0:k-1}\}}{\Pr\{y(k) | y_{1:k-1}, u_{0:k-1}\}} \\ &= \frac{\Pr\{y(k)|x(k) = \delta_N^i\} \cdot \Pr\{x(k) = \delta_N^i | y_{1:k-1}, u_{0:k-1}\}}{\sum_{j=1}^N \Pr\{y(k)|x(k) = \delta_N^j\} \cdot \Pr\{x(k) = \delta_N^j | y_{1:k-1}, u_{0:k-1}\}} \\ &:= \frac{V_i \cdot \pi_i(k|k-1)}{V^\top \pi(k|k-1)} \end{aligned}$$

where V_i means the i th entry of vector V and V is defined by the above equation, in which the event “ $y(k) = \delta_Q^{\lambda_k}$ ” is represented as “ $y(k)$ ” for simplicity. By taking the expectation operation on both sides of (4b), for the deterministic state variable $x(k) = \delta_N^i$, one has

$$\begin{aligned} \mathbb{E}\{y(k)|x(k) = \delta_N^i\} &= \mathcal{H}_k \times \mathbb{E}\{v(k)\} \times x(k) \\ &= \mathcal{H}_k \times_{i=1}^{L_2} [q_k^i \quad 1 - q_k^i]^\top \times \delta_N^i \\ &:= \text{Col}_i(H_k) \end{aligned} \quad (11)$$

where $H_k = \mathcal{H}_k \times_{i=1}^{L_2} [q_k^i \quad 1 - q_k^i]^\top$. On the other hand, it is obtained that

$$\begin{aligned} \mathbb{E}\{y(k)|x(k) = \delta_N^i\} &= \sum_{j=1}^Q \Pr\{y(k) = \delta_Q^j | x(k) = \delta_N^i\} \cdot \delta_Q^j \\ &= \begin{bmatrix} \Pr\{y(k) = \delta_Q^1 | x(k) = \delta_N^i\} \\ \vdots \\ \Pr\{y(k) = \delta_Q^Q | x(k) = \delta_N^i\} \end{bmatrix}. \end{aligned}$$

With the above in mind, we can use (11) to obtain

$$\begin{aligned} V &= \begin{bmatrix} \Pr\{y(k) = \delta_Q^{\lambda_k} | x(k) = \delta_N^1\} \\ \vdots \\ \Pr\{y(k) = \delta_Q^{\lambda_k} | x(k) = \delta_N^N\} \end{bmatrix} \\ &= \begin{bmatrix} (H_k)_{\lambda_k, 1} \\ \vdots \\ (H_k)_{\lambda_k, N} \end{bmatrix} = [\text{Row}_{\lambda_k}(H_k)]^\top \end{aligned} \quad (12)$$

and then it follows that

$$\begin{aligned} \pi_i(k|k) &= \frac{1}{V^\top \pi(k|k-1)} \begin{bmatrix} V_1 \cdot \pi_1(k|k-1) \\ \vdots \\ V_N \cdot \pi_N(k|k-1) \end{bmatrix} \\ &= \frac{V \circ \pi(k|k-1)}{V^\top \pi(k|k-1)} = \frac{[\text{Row}_{\lambda_k}(H_k)]^\top \circ \pi(k|k-1)}{[\text{Row}_{\lambda_k}(H_k)] \pi(k|k-1)} \end{aligned}$$

which completes the proof. ■

Remark 4: It is worth pointing out that the term $\text{Row}_{\lambda_k}(H_k)$ might be zero, and even if it is not zero, the denominator of (6b) might also be zero. Therefore, some numerical issues would arise when implementing the above developed algorithm. In order to avoid such kind of possible numerical problems, in the real implementation, the estimate of the forward state probability distribution vector is always modified as follows:

$$\pi(k|k) = \begin{cases} \frac{[\text{Row}_{\lambda_k}(H_k)]^\top \circ \pi(k|k-1)}{V}, & \text{if } V \neq 0 \\ \pi(k|k-1), & \text{if } V = 0 \end{cases}$$

where $\mathcal{V} \triangleq \text{Row}_{\lambda_k}(H_k)\pi(k|k-1)$.

Now, we are in the position to design the Boolean-valued state estimator $\hat{X}(k)$ by using the state probability distribution vector $\pi(k|k)$. For the convenience of later development, let us define $\hat{X}^*(k)$ as the optimal estimate of $X(k)$ based on the noisy measurements $y_{1:k}$. Let $\mathbf{S} \triangleq [X^1 \cdots X^N]$, where $X^i \in \mathcal{D}^n$ and $X^i \sim \delta_N^i$ (i.e., X^i is identified with the vector δ_N^i), $i \in \{1, 2, \dots, N\}$.

Theorem 1: Consider system (1) with noisy measurements $Y_{1:k}$ and control sequence $U_{0:k-1}$. The optimal MMSE filter $\hat{X}^*(k)$ is given by

$$\hat{X}^*(k) = \mathbf{g}(\mathbf{S}\pi(k|k)) \quad (13)$$

where the function $\mathbf{g}(\cdot) : [0, 1]^n \rightarrow \mathcal{D}^n$ is defined as $\mathbf{g}(\alpha) = (g(\alpha_1), g(\alpha_2), \dots, g(\alpha_n))^\top$ with $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^\top \in [0, 1]^n$ and $g(\cdot)$ given below:

$$g(\alpha) \triangleq \begin{cases} 1, & 0.5 \leq \alpha \leq 1 \\ 0, & 0 \leq \alpha < 0.5. \end{cases}$$

Furthermore, the corresponding optimal filtering MSE is

$$\text{MSE}[\hat{X}^*(k)] = \frac{n}{2} - \|\mathbf{S}\pi(k|k) - \Theta\|_1 \quad (14)$$

where $\Theta \in \mathbb{R}^{n \times 1}$ is a vector with all of its elements being $\frac{1}{2}$.

Proof: From the definition of the conditional MSE, one has

$$\text{MSE}[\hat{X}(k)] = \mathbb{E} \left\{ \sum_{i=1}^n (X_i(k) - \hat{X}_i(k))^2 \middle| Y_{1:k}, U_{0:k-1} \right\}. \quad (15)$$

Since $X_i(k) \in \mathcal{D}$ and $\hat{X}_i(k) \in \mathcal{D}$, it is easy to verify that

$$(X_i(k) - \hat{X}_i(k))^2 = |X_i(k) - \hat{X}_i(k)|.$$

It then follows from (15) that

$$\text{MSE}[\hat{X}(k)] = \sum_{i=1}^n \mathbb{E} \{ |X_i(k) - \hat{X}_i(k)| \middle| Y_{1:k}, U_{0:k-1} \}. \quad (16)$$

The optimal estimator $\hat{X}^*(k)$ with MMSE in (16) can be derived as

$$\hat{X}^*(k) = \arg \min_{\hat{X}(k) \in \mathcal{D}^n} \sum_{i=1}^n \mathbb{E} \{ |X_i(k) - \hat{X}_i(k)| \middle| Y_{1:k}, U_{0:k-1} \}.$$

Minimization of the above equation is derived if $\hat{X}_i(k) \in \mathcal{D}$ is chosen to minimize the expectation

$$\mathbb{E} \{ |X_i(k) - \hat{X}_i(k)| \middle| Y_{1:k}, U_{0:k-1} \}$$

for all $i = 1, \dots, n$, that is,

$$\begin{aligned} \hat{X}_i^*(k) &= \arg \min_{\hat{X}_i(k) \in \mathcal{D}} \mathbb{E} \{ |X_i(k) - \hat{X}_i(k)| \middle| Y_{1:k}, U_{0:k-1} \} \\ &= \arg \min_{\hat{X}_i(k) \in \mathcal{D}} |\mathbb{E} \{ X_i(k) \middle| Y_{1:k}, U_{0:k-1} \} - \hat{X}_i(k)| \\ &= \arg \min_{\hat{X}_i(k) \in \mathcal{D}} |\bar{X}_i(k|k) - \hat{X}_i(k)| \end{aligned} \quad (17)$$

where $\bar{X}(k|k) = (\bar{X}_1(k|k), \bar{X}_2(k|k), \dots, \bar{X}_n(k|k))^\top$ represents the mathematical expectation of state $X(k)$ conditional on measurements up to (and including) time k and the control sequence $U_{0:k-1}$. The above equation yields

$$\begin{aligned} \hat{X}_i^*(k) &= \begin{cases} 1, & \text{if } 0.5 \leq \bar{X}_i(k|k) \leq 1 \\ 0, & \text{if } 0 \leq \bar{X}_i(k|k) < 0.5 \end{cases} \\ &:= g(\bar{X}_i(k|k)) \end{aligned} \quad (18)$$

and hence

$$\hat{X}^*(k) = (\hat{X}_1^*(k), \dots, \hat{X}_n^*(k))^\top = \mathbf{g}(\bar{X}(k|k)). \quad (19)$$

Combining (17) and (18), one obtains that

$$\begin{aligned} &\mathbb{E} \{ |X_i(k) - \hat{X}_i^*(k)| \middle| Y_{1:k}, U_{0:k-1} \} \\ &= |\bar{X}_i(k|k) - g(\bar{X}_i(k|k))| \\ &= \begin{cases} 1 - \bar{X}_i(k|k), & \text{if } 0.5 \leq \bar{X}_i(k|k) \leq 1 \\ \bar{X}_i(k|k), & \text{if } 0 \leq \bar{X}_i(k|k) < 0.5 \end{cases} \\ &= 0.5 - |\bar{X}_i(k|k) - 0.5|. \end{aligned}$$

Then, the optimal conditional MSE obtained by $\hat{X}^*(k)$ can be computed as follows:

$$\begin{aligned} \text{MSE}[\hat{X}^*(k)] &= \sum_{i=1}^n (0.5 - |\bar{X}_i(k|k) - 0.5|) \\ &= \frac{n}{2} - \|\bar{X}(k|k) - \Theta\|_1. \end{aligned} \quad (20)$$

Both (19) and (20) require the computation of $\bar{X}(k|k)$, which can be obtained as

$$\begin{aligned} \bar{X}(k|k) &= \sum_{i=1}^N X^i \text{Pr} \{ X(k) = X^i \sim \delta_N^i \middle| Y_{1:k}, U_{0:k-1} \} \\ &= [X^1 \cdots X^N] \pi(k|k) := \mathbf{S}\pi(k|k). \end{aligned}$$

The proof is now complete. ■

Lemma 1 and Theorem 1 provide a recursive matrix-based method for computing the optimal estimator $\hat{X}^*(k)$ and its MMSE, whose procedure is displayed in Algorithm 1.

Algorithm 1 Boolean Bayesian Filter

Input: the noisy measurements $y_{1:k} = \{\delta_Q^{\lambda_1}, \dots, \delta_Q^{\lambda_k}\}$ and the control sequence $u_{0:k-1} = \{\delta_M^{\mu_0}, \dots, \delta_M^{\mu_{k-1}}\}$, the prediction matrix F_k , the update matrix H_k , the matrix \mathbf{S} , and the function $\mathbf{g}(\cdot)$

Output: the state probability distribution vectors $\pi(k|k)$, the Boolean-valued state estimation $\hat{X}^*(k)$, and the corresponding conditional MSE $[\text{MSE}[\hat{X}^*(k)]]$

- 1: **procedure** BOOLEAN BAYESIAN FILTER
 - 2: Initialization: Set $\pi(0|0) := \mathbb{E}\{x(0)\}$
 - 3: **for** $k = 1, 2, \dots$ **do**
 - 4: Prediction:

$$\pi(k|k-1) := (F_{k-1} \times \delta_M^{\mu_{k-1}}) \pi(k-1|k-1)$$
 - 5: Update:

$$\pi(k|k) = \begin{cases} \frac{[\text{Row}_{\lambda_k}(H_k)]^\top \circ \pi(k|k-1)}{\mathcal{V}}, & \text{if } \mathcal{V} \neq 0 \\ \pi(k|k-1), & \text{if } \mathcal{V} = 0 \end{cases}$$
 - 6: State Estimation: $\hat{X}^*(k) := \mathbf{g}(\mathbf{S}\pi(k|k))$
 - 7: Optimal Conditional MSE:

$$\text{MSE}[\hat{X}^*(k)] := \frac{n}{2} - \|\mathbf{S}\pi(k|k) - \Theta\|_1$$
 - 8: **end for**
 - 9: **return** $(\pi(k|k), \hat{X}^*(k), \text{MSE}[\hat{X}^*(k)])$
 - 10: **end procedure**
-

B. Boolean Bayesian Smoother

In the post-mortem analysis, it is often the case that some future measurements are available for the current state estimation. Based on a fixed interval of noisy measurements, *fixed-interval smoothing* method is to be developed for the state estimation of STVBN, which usually might have a better performance than the algorithm proposed in the previous section. In other words, if we have measurements $y_{1:\tau} = \{\delta_Q^{\lambda_1}, \dots, \delta_Q^{\lambda_\tau}\}$ in a fixed interval that are available, then we

desire to obtain $X(1|\tau), X(2|\tau), \dots, X(\tau|\tau)$, where $X(k|\tau)$ is the exact state value of (1) for $k = 1, 2, \dots, \tau$. Such an estimator is applicable in the case when one has recorded some measurements which are available for post-processing.

Now, we are interested in deriving an estimate of the state $X(k)$ based on the measurements from $t = 1$ to $t = \tau$ with $\tau > k$. An optimal smoothing filter is an estimator $\hat{X}(k) = \hat{g}(Y_{1:\tau}, U_{0:\tau-1})$ of the state variable $X(k)$ that minimizes the smoothing mean-square error (SMSE) defined below:

$$\text{SMSE}[\hat{X}(k)] \triangleq \mathbb{E}\{\|X(k) - \hat{X}(k)\|_2^2 | Y_{1:\tau}, U_{0:\tau-1}\}.$$

In what follows, the *forward-backward approach* will be introduced to deal with the smoothing problem. To this end, we first define the backward probability distribution vector for state $x(r)$ concerning $\overleftarrow{y}_{\tau:j} = \{\delta_Q^{\lambda_j}, \dots, \delta_Q^{\lambda_\tau}\}$ and $\overleftarrow{u}_{\tau-1:r} = \{\delta_M^{\mu_{\tau-1}}, \dots, \delta_M^{\mu_r}\}$ as

$$\overleftarrow{\pi}(r|j) \triangleq \begin{bmatrix} \Pr\{x(r) = \delta_N^1 | \overleftarrow{y}_{\tau:j}, \overleftarrow{u}_{\tau-1:r}\} \\ \vdots \\ \Pr\{x(r) = \delta_N^N | \overleftarrow{y}_{\tau:j}, \overleftarrow{u}_{\tau-1:r}\} \end{bmatrix} \quad (21)$$

for $r = \tau, \tau - 1, \dots, k$ and $j = r, r + 1$.

The smoothed probability distribution vector for state $x(k)$ is defined as

$$\Pi(k|\tau) \triangleq \begin{bmatrix} \Pr\{x(k) = \delta_N^1 | y_{1:\tau}, u_{0:\tau-1}\} \\ \vdots \\ \Pr\{x(k) = \delta_N^N | y_{1:\tau}, u_{0:\tau-1}\} \end{bmatrix}.$$

Using the forward-backward technique for smoothing, two conditional probability distribution vector estimates are obtained concerning the state $x(k)$. The first estimate, $\pi(k|k)$ (i.e., a posteriori estimate at time instant k), is based on the Boolean Bayesian filter that runs forward in time from $t = 1$ to $t = k$ by utilizing the measurements $y_{1:k}$. The second estimate, $\overleftarrow{\pi}(k|k+1)$, is based on the Boolean Bayesian filter that operates backward in time from $t = \tau$ back to $t = k$ by using the measurements $\overleftarrow{y}_{\tau:k+1}$, that is, a priori estimate at time instant k from a reversed time perspective. Then, the optimal smoothed estimate of the conditional probability distribution vector concerning the state $x(k)$, denoted by $\Pi(k|\tau)$ for convenience, can be obtained by combining these two estimates.

Note that the backward estimate needs to operate backward in time by starting at the terminal time instant τ . Since the backward and forward estimates must be independent, the information that has been utilized in the forward estimate is not allowed to be re-utilized in the backward estimate. Hence, the backward initial distribution of the state at the final time $t = \tau$ is set to be

$$\overleftarrow{\pi}(\tau|\tau+1) \triangleq \left[\frac{1}{N} \quad \frac{1}{N} \quad \dots \quad \frac{1}{N} \right]^\top. \quad (22)$$

The following result provides an efficient method for calculating the one-step prediction and estimation of the backward state probability distribution vector.

Lemma 2: Consider system (1) with algebraic form (4). Let $\overleftarrow{\pi}(\tau|\tau+1)$ be the initial distribution of the state at the terminal time instant τ . Then, one has

$$\overleftarrow{\pi}(\gamma|\gamma) = \frac{[\text{Row}_{\lambda_\gamma}(H_\gamma)]^\top \circ \overleftarrow{\pi}(\gamma|\gamma+1)}{[\text{Row}_{\lambda_\gamma}(H_\gamma)] \overleftarrow{\pi}(\gamma|\gamma+1)} \quad (23a)$$

$$\overleftarrow{\pi}(\gamma-1|\gamma) = ((\delta_M^{\mu_{\gamma-1}})^\top \times F_{\gamma-1}^\top) \Psi_{\gamma-1} \overleftarrow{\pi}(\gamma|\gamma) \quad (23b)$$

for $\gamma = \tau, \tau - 1, \dots, k + 1$, where $\overleftarrow{\pi}(\tau|\tau+1)$ is defined as in (22), and $\Psi_{\gamma-1} \triangleq \text{diag}\{1/\|\psi_{\gamma-1}^1\|_1, \dots, 1/\|\psi_{\gamma-1}^N\|_1\}$ with $\psi_{\gamma-1}^i = \text{Col}_i((\delta_M^{\mu_{\gamma-1}})^\top \times F_{\gamma-1}^\top)$, $i = 1, 2, \dots, N$.

Proof: The proof of (23a) follows a similar line as that of (6b) in Lemma 1, and is thus omitted here for the sake of brevity. In

Algorithm 2 Boolean Bayesian Smoother

Input: the noisy measurements $y_{1:\tau} = \{\delta_Q^{\lambda_1}, \dots, \delta_Q^{\lambda_\tau}\}$ and the control sequence $u_{0:\tau-1} = \{\delta_M^{\mu_0}, \dots, \delta_M^{\mu_{\tau-1}}\}$, the prediction matrix F_k , the update matrix H_k , the matrix \mathbf{S} , and the function $\mathbf{g}(\cdot)$

Output: the state probability distribution vectors $\Pi(k|\tau)$, the Boolean-valued state estimation $\hat{X}^*(k)$, and the corresponding optimal SMSE $[\hat{X}^*(k)]$

1: **procedure** BOOLEAN BAYESIAN SMOOTHER

2: Initialization: Set $\pi(0|0) := \mathbb{E}\{x(0)\}$ and

$$\overleftarrow{\pi}(\tau|\tau+1) := \left[\frac{1}{N} \quad \frac{1}{N} \quad \dots \quad \frac{1}{N} \right]^\top$$

3: **for** $\gamma := 1, 2, \dots, k$ **do**

4: Prediction:

$$\pi(\gamma|\gamma-1) := (F_{\gamma-1} \times \delta_M^{\mu_{\gamma-1}}) \pi(\gamma-1|\gamma-1)$$

5: Update:

$$\pi(k|k) = \begin{cases} \frac{[\text{Row}_{\lambda_k}(H_k)]^\top \circ \pi(k|k-1)}{\mathcal{V}}, & \text{if } \mathcal{V} \neq 0 \\ \pi(k|k-1), & \text{if } \mathcal{V} = 0 \end{cases}$$

6: **end for**

7: **for** $\gamma := \tau, \tau - 1, \dots, k + 1$ **do**

8: Update:

$$\overleftarrow{\pi}(\gamma|\gamma) := \frac{[\text{Row}_{\lambda_\gamma}(H_\gamma)]^\top \circ \overleftarrow{\pi}(\gamma|\gamma+1)}{[\text{Row}_{\lambda_\gamma}(H_\gamma)] \overleftarrow{\pi}(\gamma|\gamma+1)}$$

9: Prediction:

$$\overleftarrow{\pi}(\gamma-1|\gamma) = ((\delta_M^{\mu_{\gamma-1}})^\top \times F_{\gamma-1}^\top) \Psi_{\gamma-1} \overleftarrow{\pi}(\gamma|\gamma)$$

10: **end for**

11: Smoothed Distribution Vector:

$$\Pi(k|\tau) := \frac{\overleftarrow{\pi}(k|k+1) \circ \pi(k|k)}{\|\overleftarrow{\pi}(k|k+1) \circ \pi(k|k)\|_1}$$

12: State Estimation: $\hat{X}^*(k) := \mathbf{g}(\mathbf{S}\Pi(k|\tau))$

13: Optimal SMSE:

$$\text{SMSE}[\hat{X}^*(k)] := \frac{n}{2} - \|\mathbf{S}\Pi(k|\tau) - \Theta\|_1$$

14: **return** $(\Pi(k|\tau), \hat{X}^*(k), \text{SMSE}[\hat{X}^*(k)])$

15: **end procedure**

the following, we show the validity of (23b). Note that the matrix $F_{\gamma-1} \times \delta_M^{\mu_{\gamma-1}}$ is the transition probability matrix of the controlled Markov chain defined by the state equation (4a) [6], that is,

$$\begin{aligned} & (F_{\gamma-1} \times \delta_M^{\mu_{\gamma-1}})_{i,j} \\ &= \Pr\{x(\gamma) = \delta_N^i | x(\gamma-1) = \delta_N^j, u(\gamma-1) = \delta_M^{\mu_{\gamma-1}}\}. \end{aligned}$$

For the given $x(\gamma) = \delta_N^i$ and $u(\gamma-1) = \delta_M^{\mu_{\gamma-1}}$, the backward probability distribution vector of state $x(\gamma-1)$, denoted by $\pi(\gamma-1|x(\gamma), u(\gamma-1))$ (i.e., $\pi_j(\gamma-1|x(\gamma), u(\gamma-1)) \triangleq \Pr\{x(\gamma-1) = \delta_N^j | x(\gamma) = \delta_N^i, u(\gamma-1) = \delta_M^{\mu_{\gamma-1}}\}$ for $j = 1, 2, \dots, N$), can be obtained as

$$\begin{aligned} \pi(\gamma-1|x(\gamma), \delta_M^{\mu_{\gamma-1}}) &= \frac{[\text{Row}_i(F_{\gamma-1} \times \delta_M^{\mu_{\gamma-1}})]^\top}{\left\| [\text{Row}_i(F_{\gamma-1} \times \delta_M^{\mu_{\gamma-1}})]^\top \right\|_1} \\ &= \frac{\text{Col}_i((\delta_M^{\mu_{\gamma-1}})^\top \times F_{\gamma-1}^\top)}{\left\| \text{Col}_i((\delta_M^{\mu_{\gamma-1}})^\top \times F_{\gamma-1}^\top) \right\|_1}. \end{aligned}$$

Thus, it follows that

$$\begin{aligned}
 & \overleftarrow{\pi}(\gamma-1|\gamma) \\
 &= \sum_{i=1}^N \Pr\{x(\gamma) = \delta_N^i | \overleftarrow{y}_{\tau:\gamma}, \overleftarrow{u}_{\tau-1:\gamma}\} \cdot \pi(\gamma-1|x(\gamma), \delta_M^{\mu\gamma-1}) \\
 &= \sum_{i=1}^N \overleftarrow{\pi}_i(\gamma|\gamma) \frac{\text{Col}_i((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top)}{\|\text{Col}_i((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top)\|_1} \\
 &= \left[\frac{\text{Col}_1((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top)}{\|\text{Col}_1((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top)\|_1} \right. \\
 &\quad \left. \cdots \frac{\text{Col}_N((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top)}{\|\text{Col}_N((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top)\|_1} \right] \overleftarrow{\pi}(\gamma|\gamma) \\
 &= ((\delta_M^{\mu\gamma-1})^\top \times F_{\gamma-1}^\top) \Psi_{\gamma-1} \overleftarrow{\pi}(\gamma|\gamma),
 \end{aligned}$$

which completes the proof. \blacksquare

Lemma 3: Consider system (1) with noisy measurements $y_{1:\tau}$ and control sequence $u_{0:\tau-1}$. The smoothed conditional probability distribution vector $\Pi(k|\tau)$ is derived as

$$\Pi(k|\tau) = \frac{\overleftarrow{\pi}(k|k+1) \circ \pi(k|k)}{\|\overleftarrow{\pi}(k|k+1) \circ \pi(k|k)\|_1}. \quad (24)$$

Proof: Based on the Bayes' theorem, one can verify that (25) holds which is shown at the bottom of the next page. Notice that

$$\begin{aligned}
 & \Pr\{\overleftarrow{y}_{\tau:k+1}, \overleftarrow{u}_{\tau-1:k} | x(k) = \delta_N^i\} \\
 &= \frac{\Pr\{x(k) = \delta_N^i | \overleftarrow{y}_{\tau:k+1}, \overleftarrow{u}_{\tau-1:k}\} \cdot \Pr\{\overleftarrow{y}_{\tau:k+1}, \overleftarrow{u}_{\tau-1:k}\}}{\Pr\{x(k) = \delta_N^i\}}
 \end{aligned}$$

which, together with (25), implies that (24) holds. \blacksquare

With the help of Theorem 1 and Lemma 3, we have the following result immediately.

Theorem 2: Consider system (1) with noisy measurements $y_{1:\tau}$ and control sequence $u_{0:\tau-1}$. The optimal minimum SMSE filter $\hat{X}^*(k)$ is designed as follows

$$\hat{X}^*(k) = \mathbf{g}(\mathbf{S}\Pi(k|\tau)) \quad (26)$$

and the corresponding optimal SMSE can be calculated by

$$\text{SMSE}[\hat{X}^*(k)] = \frac{n}{2} - \|\mathbf{S}\Pi(k|\tau) - \Theta\|_1.$$

The pseudocode for computing the optimal smoother and its minimum SMSE is displayed in Algorithm 2.

Remark 5: The relating issues concerning the state estimation of BNs have been discussed recently in [15], where both the minimum mismatching estimation and the maximum posterior estimation of the state have been provided for BCNs with stochastic disturbance. When referring to the state estimation for STVBNs, the similar discussions are much more complex. Moreover, it is worth pointing out that in [15], the conditional probability distribution of the state is calculated based on only the historical control inputs and measurement outputs. However, in the post-mortem analysis, it is often the case that some future measurements are available for the state estimation. Based on a fixed interval of noisy measurements, the *forward-backward approach* is utilized for the state estimation of STVBNs, which usually might have a better performance than the algorithms proposed in [15].

Remark 6: The main results in this paper (i.e., Theorems 1-2) establish a general theoretical framework for the state estimation problem of stochastic time-varying BNs, and the developed matrix-based algorithms (i.e., Algorithms 1-2) provide an efficient way to compute the state estimation with minimum MSE for the STVBNs. However, F_k and H_k involved in the above mentioned algorithms are respectively $N \times NM$ and $Q \times N$ matrices, which means that the dimensions of F_k and H_k will grow exponentially as the size of

the corresponding network increases. Hence, the designed algorithms are applicable only to STVBNs with scales not too large.

Remark 7: In Theorems 1-2, a general theoretical framework is established for the state estimation problem of STVBNs. Our main results distinguish with the existing ones in the following two aspects: 1) the one-step prediction and estimation of the state probability distribution vector are calculated by using a new recursive matrix-based algorithm with the help from STP technique and Bayes' theorem, thereby facilitating the design of the desired Boolean Bayesian filter; and 2) the fixed-interval smoothing filter is designed using the forward-backward approach.

IV. A BIOLOGICAL EXAMPLE

The well-known *p53-MDM2* negative-feedback GRN contains a single Boolean signal Z (corresponding to the DNA damage input) and four Boolean state variables X_1, \dots, X_4 (representing the activity/inactivity of four different genes: *ATM*, *p53*, *Wip1*, *Mdm2*, respectively). In the recent paper [21], a Boolean model has been established to describe the regulation mechanism of *p53-MDM2* negative-feedback GRN as follows:

$$\begin{cases} X_1(k) = \overline{X_3(k-1)} \wedge (X_1(k-1) \vee Z(k-1)), \\ X_2(k) = \overline{X_4(k-1)} \wedge (X_1(k-1) \vee X_3(k-1)), \\ X_3(k) = X_2(k-1), \\ X_4(k) = \overline{X_1(k-1)} \wedge (X_2(k-1) \vee X_3(k-1)) \end{cases} \quad (27)$$

where " \wedge " and " \vee " are the logical operators "AND" and "OR", respectively.

With the help of STP and setting $x(t) \triangleq \times_{i=1}^4 x_i(t)$, the BN model (27) can be converted into its algebraic form as follows:

$$x(k) = Fz(k-1)x(k-1) \quad (28)$$

where $z(k-1)$ is the vector form of the Boolean signal $Z(k-1)$ and $F \in \mathcal{L}_{16 \times 32}$.

The BN model (27) has two different contexts, where the value of the external signal z determines which context is active. Specifically, the input z in (28) may attain one of its possible two values: DNA damage $z = \delta_1^1$ or no-stress $z = \delta_2^2$. For each of these two possible values, $F_i \triangleq F \times \delta_i^i$ ($i = 1, 2$) is a 16×16 logical matrix that can be treated as the state transition matrix of the corresponding constituent BN, denoted by BN_i for simplicity.

Now, let us discuss the context estimation problem of BN model (27) based on the noisy measurements. As a matter of fact, the unobserved context that selects between the constituent BNs can be encoded by the external signal Z , that is, $Z = 1$ for BN_1 and $Z = 0$ for BN_2 . The evolution of the context process is assumed to be independent with the other state variables. The logical equation for the context selector Z is governed by

$$Z(k) = Z(k-1) \oplus W(k-1) \quad (29)$$

where $W(k) \in \mathcal{D}$ is the stochastic Bernoulli-type white noise with $W(k) \sim B(1, p_k)$, the logical operator " \oplus " leaves the state of the context selector $Z(k)$ unaltered when $W(k-1) = 0$ and flips it when $W(k-1) = 1$. The parameter p_k indicates the intensity of the noise, which can be regarded as the context switching probability of system (27) determining the possibility that the context will switch between BN_1 and BN_2 at a given time instant k .

To facilitate the context estimation problem, we put together the state and the context processes to obtain the following augmented system:

$$\begin{cases} X(k) = f(X(k-1), Z(k-1)) \\ Z(k) = Z(k-1) \oplus W(k-1) \end{cases} \quad (30a)$$

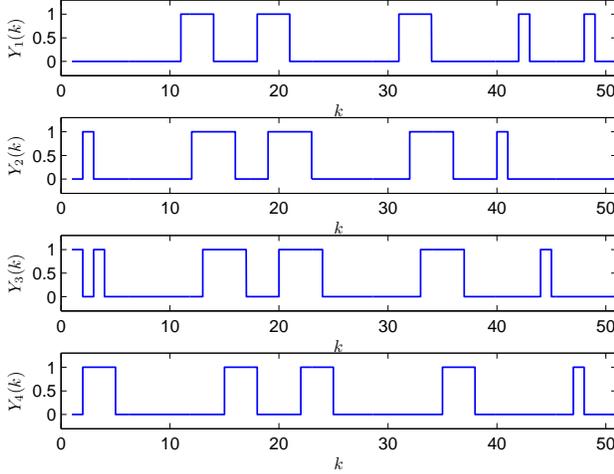


Fig. 1. The real observation value of the measurement output in (30b).

where $f(\cdot, \cdot)$ is defined the same as in (27). The measurement output in this example is given as

$$Y(k) = X(k) \oplus V(k) \quad (30b)$$

which implies that only the state variables X_1, \dots, X_4 are observable under noisy conditions. Here, the measurement noise $V(k) = (V_1(k), \dots, V_4(k))$ is assumed to be white with $V_j(k) \sim B(1, q_k^j)$ for $j = 1, \dots, 4$.

In this example, we take $p_k \equiv 0.08$ and $q_k^j \equiv 0.01$. Fig. 1 displays the real observation value of the measurement output $Y(k)$, from which we can see that, under no damage, the system spends a significant amount of time in the equilibrium state $(0, 0, 0, 0)$, whereas under DNA damage, more states are visited periodically. Fig. 2 presents the value of the conditional expectation $\bar{Z}(k|k) = \mathbb{E}\{Z(k)|Y_{1:k}\}$ and the optimal MMSE filter $\hat{Z}^*(k)$ returned by Algorithm 1 as well as the context selector $Z(k)$. It can be observed that the filter $\hat{Z}^*(k)$ performs quite well in tracking the true signal $Z(k)$. From Theorem 1, one obtains

$$\text{MSE}[\hat{Z}^*(k)] = \frac{1}{2} - \left| \bar{Z}(k|k) - \frac{1}{2} \right|$$

which implies that $|\bar{Z}(k|k) - \frac{1}{2}|$ can be regarded as the *confidence* in the estimation of the context [4]. We can see clearly from Fig. 2 that the confidence become small near the transition of the context. Moreover, there always exists a lag before $\hat{Z}^*(k)$ responds to a transition in $Z(k)$ (e.g., the transition at time instants $k = 11, 23, 31, 38$). During this lag, enough contradictory measurements are accumulated in order to switch the estimation of the context.

In addition, Fig. 3 depicts the MSE returned by Algorithm 1 and the SMSE calculated by Algorithm 2, from which it can be observed

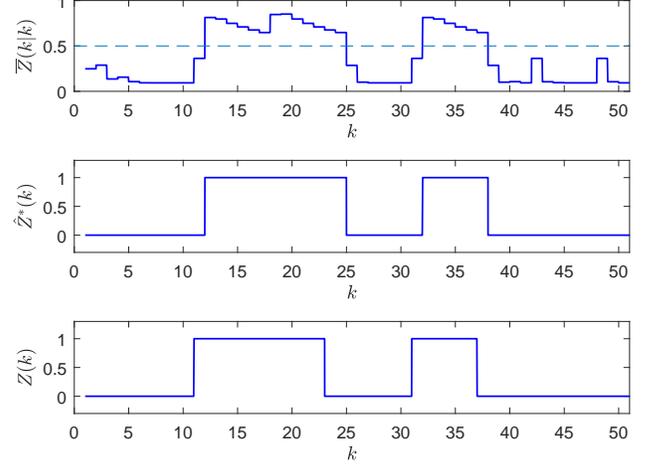


Fig. 2. Estimated and true context trajectories of BN model (27).

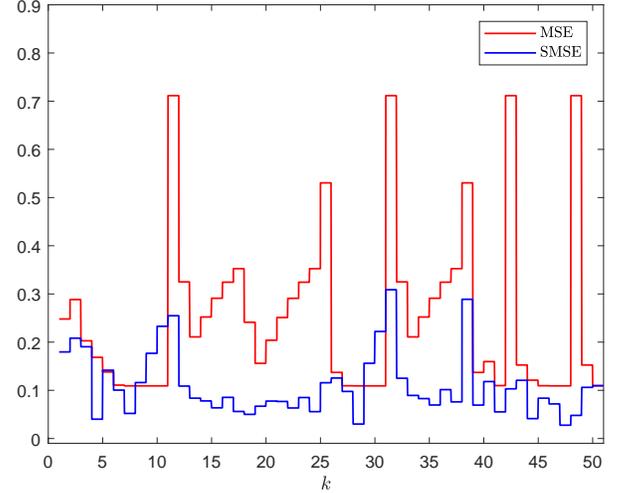


Fig. 3. MSE and SMSE returned by the proposed algorithms.

that the SMSE is much smaller than the MSE in most of the time instants, since the Boolean Bayesian smoother also uses the future measurements, whereas the Boolean Bayesian filter uses only the measurements up to the present time instant. It is worth pointing out that, in some particular/extreme situations, the performance of the Boolean Bayesian smoother might not be better than that of the Boolean Bayesian filter as expected. Fortunately, such cases usually do not occur.

$$\begin{aligned} \Pr\{x(k) = \delta_N^i | y_{1:\tau}, u_{0:\tau-1}\} &= \frac{\Pr\{\bar{y}_{\tau:k+1}, \bar{u}_{\tau-1:k} | x(k) = \delta_N^i, y_{1:k}, u_{0:k-1}\} \cdot \Pr\{x(k) = \delta_N^i | y_{1:k}, u_{0:k-1}\}}{\Pr\{\bar{y}_{\tau:k+1}, \bar{u}_{\tau-1:k} | y_{1:k}, u_{0:k-1}\}} \\ &= \frac{\Pr\{\bar{y}_{\tau:k+1}, \bar{u}_{\tau-1:k} | x(k) = \delta_N^i\} \cdot \Pr\{x(k) = \delta_N^i | y_{1:k}, u_{0:k-1}\}}{\sum_{j=1}^N \Pr\{x(k) = \delta_N^j, \bar{y}_{\tau:k+1}, \bar{u}_{\tau-1:k} | y_{1:k}, u_{0:k-1}\}} \\ &= \frac{\Pr\{\bar{y}_{\tau:k+1}, \bar{u}_{\tau-1:k} | x(k) = \delta_N^i\} \cdot \Pr\{x(k) = \delta_N^i | y_{1:k}, u_{0:k-1}\}}{\sum_{j=1}^N \Pr\{\bar{y}_{\tau:k+1}, \bar{u}_{\tau-1:k} | x(k) = \delta_N^j\} \cdot \Pr\{x(k) = \delta_N^j | y_{1:k}, u_{0:k-1}\}} \end{aligned} \quad (25)$$

V. CONCLUSION

In this paper, the state estimation problem has been investigated for a class of time-varying BNs subjected to stochastic disturbances, which have been simulated by a series of mutually independent random variables obeying Bernoulli distribution with certain success probabilities. An algebraic representation of the STVBNs has been obtained by resorting to the STP technique. Consequently, Boolean Bayesian filter has been developed to estimate the true states of the STVBNs, and a recursive matrix-based algorithm has been put forward to calculate the one-step prediction and estimation of the state probability distribution vector. Moreover, the fixed-interval smoothing filter has also been achieved by applying the forward-backward approach. Finally, a simulation experiment has been employed to illustrate the applicability and effectiveness of the proposed algorithms.

For the derived theoretical results, their limitation lies in the computational complexity. Note that the state space grows exponentially as the size of the corresponding network increases, the designed filters in this paper are available only to STVBNs with small scales. In the near future, we will focus on the topics of decreasing the computational complexities concerning STVBNs so as to improve the established filtering algorithms.

REFERENCES

- [1] T. Akutsu, M. Hayashida, W.-K. Ching, and M. K. Ng, "Control of Boolean networks: Hardness results and algorithms for tree structured networks," *J. Theor. Biol.*, vol. 244, no. 4, pp. 670–679, Feb. 2007.
- [2] A. Bahadorinejad and U. M. Braga-Neto, "Optimal fault detection and diagnosis in transcriptional circuits using next-generation sequencing," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 15, no. 2, pp. 516–525, Mar./Apr. 2007.
- [3] N. Bof, E. Fornasini, and M. E. Valcher, "Output feedback stabilization of Boolean control networks," *Automatica*, vol. 57, pp. 21–28, Jul. 2015.
- [4] U. M. Braga-Neto, "Optimal state estimation for Boolean dynamical systems," in *Proc. 45th Annu. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, 2011, pp. 1050–1054.
- [5] T. Chen and U. M. Braga-Neto, "Maximum-likelihood estimation of the discrete coefficient of determination in stochastic Boolean systems," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3880–3894, Aug. 2013.
- [6] H. Chen and J. Liang, "Local Synchronization of interconnected Boolean networks with stochastic disturbances," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 2, pp. 452–463, Feb. 2020.
- [7] D. Cheng, "Disturbance decoupling of Boolean control networks," *IEEE Trans. Autom. Control*, vol. 56, no. 1, pp. 2–10, Jan. 2011.
- [8] D. Cheng, T. Liu, K. Zhang, and H. Qi, "On decomposed subspaces of finite games," *IEEE Trans. Autom. Control*, vol. 61, no. 11, pp. 3651–3656, Nov. 2016.
- [9] D. Cheng, H. Qi, and Z. Li, *Analysis and Control of Boolean Networks: A Semi-tensor Product Approach*. London, U.K.: Springer-Verlag, 2011.
- [10] D. Cheng, H. Qi, Z. Li, and J. B. Liu, "Stability and stabilization of Boolean networks," *Int. J. Robust Nonlinear Control*, vol. 21, no. 2, pp. 134–156, Jan. 2011.
- [11] A. Fauré and D. Thieffry, "Logical modelling of cell cycle control in eukaryotes: A comparative study," *Mol. Biosyst.*, vol. 5, no. 12, pp. 1569–1581, 2009.
- [12] E. Fornasini and M. E. Valcher, "Optimal control of Boolean control networks," *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1258–1270, May 2014.
- [13] E. Fornasini and M. E. Valcher, "Fault detection analysis of Boolean control networks," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2734–2739, Oct. 2015.
- [14] A. Fujita, J. R. Sato, H. M. Garay-Malpartida, P. A. Morettin, M. C. Sogayar, and C. E. Ferreira, "Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method," *Bioinformatics*, vol. 23, no. 13, pp. 1623–1630, Jul. 2007.
- [15] Y. Guo, Q. Li, and W. Gui, "Optimal state estimation of Boolean control networks with stochastic disturbances," *IEEE Trans. Cybern.*, vol. 50, no. 3, pp. 1355–1359, Mar. 2020.
- [16] M. Imani and U. M. Braga-Neto, "Maximum-likelihood adaptive filter for partially observed Boolean dynamical systems," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 359–371, Jan. 2017.
- [17] M. Imani and U. M. Braga-Neto, "Particle filters for partially-observed Boolean dynamical systems," *Automatica*, vol. 87, pp. 238–250, Jan. 2018.
- [18] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," *J. Theor. Biol.*, vol. 22, no. 3, pp. 437–467, 1969.
- [19] D. Laschov and M. Margaliot, "A maximum principle for single-input Boolean control networks," *IEEE Trans. Autom. Control*, vol. 56, no. 4, pp. 913–917, Apr. 2011.
- [20] D. Laschov and M. Margaliot, "Minimum-time control of Boolean networks," *SIAM J. Control Optim.*, vol. 5, no. 4, pp. 2869–2892, 2013.
- [21] R. K. Layek, A. Datta, and E. R. Dougherty, "From biological pathways to regulatory networks," *Mol. BioSyst.*, vol. 7, no. 3, pp. 843–851, 2011.
- [22] F. Li, "Robust stabilization for a logical system," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 6, pp. 2176–2184, Nov. 2017.
- [23] F. Li and Y. Tang, "Pinning control design for the stabilization of Boolean networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 7, pp. 1585–1590, Jul. 2016.
- [24] R. Li and T. Chu, "Complete synchronization of Boolean networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 5, pp. 840–846, May 2012.
- [25] H. Li, L. Xie, and Y. Wang, "On robust control invariance of Boolean control networks," *Automatica*, vol. 68, pp. 392–396, Jun. 2016.
- [26] J. Liang, H. Chen, and J. Lam, "An improved criterion for controllability of Boolean control networks," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 6012–6018, Nov. 2017.
- [27] J. Lu, J. Zhong, L. Li, D. W. C. Ho, and J. Cao, "Synchronization analysis of master-slave probabilistic Boolean networks," *Sci. Rep.*, vol. 5, p. 13437, Aug. 2015.
- [28] M. Meng, J. Lam, J.-e. Feng, and X. Li, " l_1 -gain analysis and model reduction problem for Boolean control networks," *Inf. Sci.*, vol. 348, pp. 68–83, Jun. 2016.
- [29] M. Meng, L. Liu, and G. Feng, "Stability and l_1 gain analysis of Boolean networks with Markovian jump parameters," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 4222–4228, Aug. 2017.
- [30] J. Paulsson, "Summing up the noise in gene networks," *Nature*, vol. 427, no. 6973, pp. 415–418, Jan. 2004.
- [31] R. Thomas, "Boolean formalization of genetic control circuits," *J. Theor. Biol.*, vol. 42, no. 3, pp. 563–585, 1973.
- [32] Z. Wang, X. Liu, Y. Liu, J. Liang and V. Vinciotti, "An extended Kalman filtering approach to modelling nonlinear dynamic gene regulatory networks via short gene expression time series," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 6, no. 3, pp. 410–419, Jul.-Sept. 2009.
- [33] Z. Wang, J. Lam, G. Wei, K. Fraser and X. Liu, "Filtering for nonlinear genetic regulatory networks with stochastic disturbances," *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2448–2457, Nov. 2008.
- [34] Z. Wang, F. Yang, D. W. C. Ho, S. Swift, A. Tucker and X. Liu, "Stochastic dynamic modeling of short gene expression time series data," *IEEE Transactions on NanoBioscience*, vol. 7, no. 1, pp. 44–55, Mar. 2008.
- [35] Y. Wu and T. Shen, "A finite convergence criterion for the discounted optimal control of stochastic logical networks," *IEEE Trans. Autom. Control*, vol. 63, no. 1, pp. 262–268, Jan. 2018.
- [36] X. Xu and Y. Hong, "Matrix approach to model matching of asynchronous sequential machines," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2974–2979, Nov. 2013.
- [37] K. Zhang and L. Zhang, "Controllability of probabilistic Boolean control networks with time-variant delays in states," *Sci. China-Inf. Sci.*, vol. 59, no. 9, p. 1092204, Sep. 2016.
- [38] K. Zhang and L. Zhang, "Observability of Boolean control networks: a unified approach based on finite automata," *IEEE Trans. Autom. Control*, vol. 61, no. 9, pp. 2733–2738, Sep. 2016.
- [39] J. Zhong, B. Li, Y. Liu, and W. Gui, "Output feedback stabilizer design of Boolean networks based on network structure," *Front. Inform. Technol. Elect. Eng.*, to be published, doi: 10.1631/FITEE.1900229.
- [40] S. Zhu, J. Lu, and D. W. C. Ho, "Finite-time stability of probabilistic logical networks: a topological sorting approach," *IEEE Trans. Circuits Syst. II-Express Briefs*, to be published, doi: 10.1109/TCSII.2019.2919018.
- [41] Y. Zou and J. Zhu, "System decomposition with respect to inputs for Boolean control networks," *Automatica*, vol. 50, no. 4, pp. 1304–1309, Apr. 2014.