Optimal State Estimation for Boolean Dynamical Systems

Ulisses Braga-Neto
Department of Electrical Engineering
Texas A & M University
College Station, Texas 77843
Email: ub@ieee.org

***Abstract*—A novel state-space signal model is proposed for discrete-time Boolean dynamical systems. State evolution is governed by Boolean functions plus binary noise. The current system state is observed through an arbitrary function plus observation noise. The optimal recursive MMSE estimator for this model is called the Boolean Kalman filter (BKF), and an efficient algorithm is presented for its exact computation. The BKF is illustrated through an example of optimal context inference for Probabilistic Boolean Networks.**

***Index Terms*—CoD estimation, maximum liklihood estimation, stochastic logic model, Boolean network with perturbation.**

I. Introduction

Many modern scientific applications involve complex dynamical systems based on switching bistable components, i.e., Boolean switches, the relationship among which is described by networks of logical gates updated and observed at discrete time intervals. This is the case, notably, of many biochemical systems, such as DNA regulatory circuits [1]. Examples in technology abound, including digital logic systems, digital communication systems, and more. In fact, N. Wiener famously wrote in *Cybernetics*, more than six decades ago, about "the essential unity of the set of problems centering about communication, control, and statistical mechanics, whether in the machine or in living tissue" [2].

The present paper introduces a novel signal model for Boolean dynamic systems. The model consists of a Boolean state process partially observed through noise. State evolution is governed by Boolean functions (i.e., logic gates) plus binary noise, whereas the observation process is quite general, consisting of an arbitrary function of the state and additive noise; the observation process may therefore take on discrete or continuous values. The optimal recursive MMSE estimator for this model is called the Boolean Kalman filter (BKF). An efficient algorithm for its exact computation is provided here.

The proposed model belongs to the class of Hidden Markov Models (HMM) [3], where the "hidden" state process constitutes a Markov Chain, which is accessed indirectly via the observational process. The novelty in the proposed signal model is that it describes the relationships among the Boolean state variables through functional relationships (logical gates plus noise), rather than general transition probabilities. The proposed model unifies and generalizes several distinct models in widespread use, such as the Boolean Network (BN) model [1] the Boolean Network with perturbation (BNp) model [4], and the Probabilistic Boolean Network (PBN) model [5].

This paper is organized as follows. Section 2 reviews the Boolean Network (BN) model through an example. Section 3 introduces the proposed signal model and Section 4 gives its optimal (MMSE) state estimator, the Boolean Kalman Filter. Section 5 illustrates the application of the BKF in the optimal estimation of the context of Probabilistic Boolean Network (PBN) model. Finally, Section presents conclusions and comments on extensions. Theorems are given without proof due to space constraints.

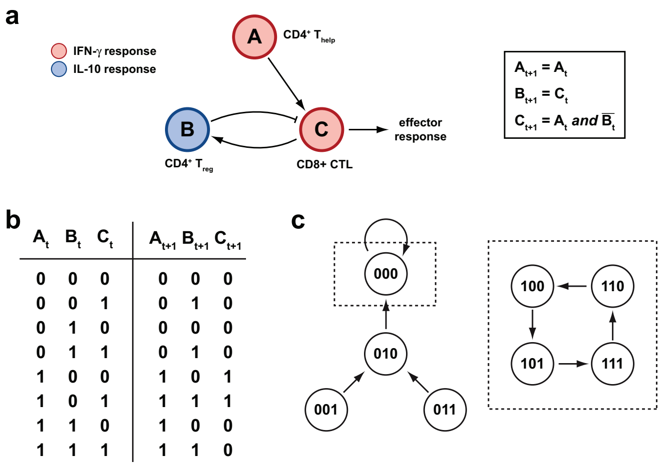
II. Boolean Networks

The Boolean Network (BN) model [1] consists of a deterministic vector time series $\{\mathbf{X}_k; k = 0, 1, \ldots\}$ specified by

$$\mathbf{X}_k = \mathbf{f}(\mathbf{X}_{k-1}), \tag{1}$$

for $k = 1, 2, \ldots$, where $\mathbf{f} : \{0,1\}^d \to \{0,1\}^d$ is an arbitrary *network function*. The latter can be written in terms of its components, $\mathbf{f} = (f_1, f_2, \ldots, f_d)$, where each component $f_i : \{0,1\}^d \to \{0,1\}$, $i = 1, \ldots, d$, is a *Boolean function*, which expresses a logical relationship among the system variables, i.e., the components $\mathbf{X}(i)$ of vector $\mathbf{X}$, for $i = 1, \ldots, d$. Due to the finite number of variables and states, the system must eventually reach recurring states, called *attractors*. These states may be singletons or part of *attractor cycles*. Every state leads to a unique attractor, so that the state space is partitioned into *attractor basins*, where each basin contains all the states that "flow" into a given attractor.

These notions are illustrated with a simple Boolean regulatory network model for immune system response [6]; see Figure 1. This is a simplified, minimal model of the immune system during infection, which nevertheless is able to capture relevant immunological behavior. The model consists of three state variables, which represent immune activation ("ON" and "OFF") of three distinct populations of T cells. Figure 1(a) depicts the three nodes, labeled "A", "B", and "C", and their activation graph. The "cytotoxic" effector response occurs when node "C" is "ON." We can see that "A" has a promoting effect on "C", while "B" inhibits it. The corresponding Boolean functions that govern state transition are displayed on the right. Figure 1(b) gives the *truth table* that determines function $\mathbf{f}$ in (1). Figure 1(c) depicts the resulting state-space.



A_t	B_t	C_t	A_{t+1}	B_{t+1}	C_{t+1}
0	0	0	0	0	0
0	0	1	0	1	0
0	1	0	0	0	0
0	1	1	0	1	0
1	0	0	1	0	1
1	0	1	1	1	1
1	1	0	1	0	0
1	1	1	1	1	0

Fig. 1. Example of a simple Boolean network model of immune system response, consisting of three nodes A, B, and C; node A is a promoter, node B is a suppressor, while node C produces the effector response, while also promoting suppression of B (negative feedback). **a**. Network wiring diagram and transition rules. **b**. State transition truth table. **c**. Basins of attraction in state-space, with attractors indicated by dashed rectangles. From Braga-Neto and Marques, "From Functional Genomics to Functional Immunomics: New Challenges, Old Problems, Big Rewards." [6].

There are only three nodes, and thus only 8 possible states. We can see that these states are partitioned into two basins of attraction: the first one corresponds to a single attractor, whereas the second one consists of an attractor cycle. As can be seen, which of the two behaviors the system is in depends only upon the state of node "A." If it is off (there is no "helper T-cell" response), then the system may pass through some irrelevant transient states but it will always tend to the resting single-state attractor 000 and thus an absence of activity; see diagram on the left in Figure 1(c). If "A" is "ON" (there is help), then the activity of the system corresponds to that of a attractor cycle with the effector response being turned on and off cyclically; see diagram on the right in Figure 1(c). This situation corresponds to modulation of the effector response and regulation of the inflammatory response by means of a negative feedback mechanism. Note that when there ceases to be a response to epitope A, the system jumps to the other basin of attraction, and tends to the resting 000 state.

III. Boolean Dynamical System Model

We propose a novel signal model that extends the class of Boolean Network models in two ways: (1) it allows for uncertainty (noise) in state transitions; (2) it allows for partial observation of the state variables and for observation noise.

A. State Model

We assume that the system is described by a *state process* $\{\mathbf{X}_k; k = 0, 1, \ldots\}$, where $\mathbf{X}_k \in \{0,1\}^d$ is a Boolean vector of size d. The evolution of the state process is governed by the equation:

$$\mathbf{X}_k = \mathbf{f}(\mathbf{X}_{k-1}) \oplus \mathbf{n}_k , \tag{2}$$

for $k = 1, 2, \ldots$, where "$\oplus$" indicates component-wise modulo-2 addition, $\mathbf{f} : \{0,1\}^d \to \{0,1\}^d$ is a network function (see the previous Section), whereas $\{\mathbf{n}_k; k = 1, 2, \ldots\}$ is a white noise process, with $\mathbf{n}_k \in \{0,1\}^d$. The noise process is "white" in the sense that $\mathbf{n}_k$ is an independent process, which is independent from the initial state $\mathbf{X}_0$; its distribution is otherwise arbitrary. The *state equation* (2) implies that the state process is Markovian:

$$\begin{aligned} P(\mathbf{X}_k = \mathbf{x}_k \mid \mathbf{X}_{k-1} = \mathbf{x}_{k-1}, \ldots, \mathbf{X}_0 = \mathbf{x}_0) \\ = P(\mathbf{X}_k = \mathbf{x}_k \mid \mathbf{X}_{k-1} = \mathbf{x}_{k-1}) \\ = P(\mathbf{n}_k = \mathbf{x}_k \ominus \mathbf{f}(\mathbf{x}_{k-1})) \end{aligned} \tag{3}$$

where "$\ominus$" denotes modulo-2 subtraction (this notation is introduced for clarity and symmetry, but it is clear that in modulo-2 arithmetics the addition and subtraction operations coincide).

The state equation (2) differs from (1) by the presence of the additive noise process. It therefore extends the previous Boolean Network model to allow for stochasticity. In the special case where the noise vector is i.i.d., with $P(\mathbf{n}_k(i) = 1) = p$ for $i = 1, \ldots, d$, then there is a probability p that each state variable $\mathbf{X}_k(i)$ will be flipped from 0 to 1 or 1 to 0, independently of the other state variables. The parameter p determines the intensity of the noise, i.e., how often a state variable is flipped. If p is very small, the system state evolves as a standard Boolean Network, settling into attractors in the long range, but the occasional flip of a variable can pull the system out of its attractor states, and even into different attractor basins altogether. On the other hand, larger p leads to much more chaotic behavior. Such a model has been known in the Systems Biology literature as a *Boolean Network with perturbation* (BNp), with parameter p [4]. The proposed state model thus generalizes both the BN and the BNp models.

B. Observation Model

The second component of the proposed signal model is the observational model. In most real-world applications, the system state is only partially observable, and distortion is introduced in the observations by environmental or sensor noise. An example is furnished by mass spectrometry measurements for proteomic applications, where the system state (protein abundances) must be inferred from noisy observations of peptides (protein fragments) [7].

Let $\mathbf{Y}_k$ be the observation corresponding to the state $\mathbf{X}_k$ at time k. Let O be the *observation space*, i.e., the set where Y_k takes its values. This defines an *observation process* $\{\mathbf{Y}_k; k = 1, 2, \ldots\}$, with $\mathbf{Y}_k \in O$ (note that the there is no observation for the initial state $\mathbf{X}_0$). The observation $\mathbf{Y}_k$ is formed from the state $\mathbf{X}_k$ through the equation:

$$\mathbf{Y}_k = \mathbf{h}(\mathbf{X}_k) \oplus \mathbf{v}_k , \tag{4}$$

for $k = 1, 2, \ldots$, where "$\oplus$" denotes an addition operator, the nature of which depends on the observation space O. For example, if the observations are real-valued variables, $O = R^m$, then "$\oplus$" stands for the usual vector-space addition, whereas if the observations are Boolean variables, $O = \{0,1\}^m$, then "$\oplus$" stands for component-wise modulo-2 addition. Both cases, and more, can be handled using essentially the same approach. The observation noise process $\{\mathbf{v}_k; k = 1, 2, \ldots\}$, with $\mathbf{v}_k \in O$, is "white" in the sense that it is an independent process, which is also independent from the state process $\{\mathbf{X}_k; k = 0, 1, \ldots\}$ and the state noise process $\{\mathbf{n}_k; k = 1, 2, \ldots\}$.

C. Signal Model

The proposed signal model puts together the state and observation equations described previously:

$$\mathbf{X}_k = \mathbf{f}(\mathbf{X}_{k-1}) \oplus \mathbf{n}_k \quad \text{(state model)}$$

$$\mathbf{Y}_k = \mathbf{h}(\mathbf{X}_k) \diamondplus \mathbf{v}_k \quad \text{(observation model)}$$

The proposed signal model describes the logical relationship between the Boolean state variables, while allowing stochastic noise in the state transitions— generalizing thus the BN and BNp models, as mentioned previously. Furthermore, the proposed model allows for incomplete observations of the system state, which leads to the inclusion of another popular Boolean model, the Probabilistic Boolean Network model, as a special case (see Section V).

Due to the Boolean nature of the state model, the MMSE estimator of the state process is called here the *Boolean Kalman Filter* (BKF). The recursive algorithm for its exact computation is given in the next Section.

IV. Boolean Kalman Filter

The optimal filtering problem consists of finding an estimator of the state $\mathbf{X}_k$ given the previous observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$, i.e. to find a function $\hat{\mathbf{X}}_k = h(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$, taking values in $\{0,1\}^d$, which optimizes a given performance criterion among all possible functions of $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$. Consider the following two criteria, namely, the conditional mean-square error (MSE):

$$\text{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) = E\left[||\hat{\mathbf{X}}_k - \mathbf{X}_k||^2 \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right] \quad (5)$$

and the (unconditional) mean-square error

$$\text{MSE} = E\left[||\hat{\mathbf{X}}_k - \mathbf{X}_k||^2\right] = E\left[\,\text{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)\,\right]. \quad (6)$$

An estimator that minimizes (5) is optimal for the given history of observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$ (which might be all that is required). On the other hand, an estimator that minimizes (6) is optimal, in an average sense, over all possible histories $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$.

Given a vector $\mathbf{v} \in R^d$, define $\overline{\mathbf{v}} \in \{0,1\}^d$ by $\overline{\mathbf{v}}(i) = I_{\mathbf{v}(i)>1/2}$ for $i = 1, \ldots, d$, i.e. $\overline{\mathbf{v}}$ is given by component-wise thresholding at $1/2$. In addition, define

$$\overline{E}\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right] = \overline{E\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right]}. \quad (7)$$

We have the following result about the MMSE estimator of the state given the observations.

Theorem 1. *The optimal minimum MSE estimator $\hat{\mathbf{X}}_k$ of the state $\mathbf{X}_k$ given the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$ up to time k, according to either criterion (5) or (6), is given by*

$$\hat{\mathbf{X}}_k = \overline{E}\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right] \quad (8)$$

Furthermore, the optimal conditional MSE is given by

$$\begin{aligned}&\text{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) = \\ &\|\min\{E\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right], E\left[{\mathbf{X}_k}^c \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right]\}\|_1\end{aligned} \quad (9)$$

where $\|\cdot\|_1$ denotes the L_1-norm, the minimum is applied component-wise, and ${\mathbf{X}_k}^c$ denotes the complement of $\mathbf{X}_k$, given by ${\mathbf{X}_k}^c(i) = 1 - \mathbf{X}_k(i)$, for $i = 1, \ldots, d$.

The optimal solution in the previous theorem can be computed very efficiently in a recursive fashion, as will be shown next. Let $(\mathbf{x}^1, \ldots, \mathbf{x}^{2^d})$ be an arbitrary enumeration of the possible state vectors. For each time $k = 1, 2, \ldots$ define the posterior distribution vectors (PDV) $\mathbf{\Pi}_{k|k}$ and $\mathbf{\Pi}_{k|k-1}$ of length 2^d by means of

$$\mathbf{\Pi}_{k|k}(i) = P\left(\mathbf{X}_k = \mathbf{x}^i \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right), \quad (10)$$

$$\mathbf{\Pi}_{k|k-1}(i) = P\left(\mathbf{X}_k = \mathbf{x}^i \mid \mathbf{Y}_{k-1}, \ldots, \mathbf{Y}_1\right), \quad (11)$$

for $i = 1, \ldots, 2^d$. Let the *prediction matrix* M_k of size $2^d \times 2^d$ be the transition matrix of the Markov chain defined by the state model (2):

$$\begin{aligned}(M_k)_{ij} &= P(\mathbf{X}_k = \mathbf{x}^i \mid \mathbf{X}_{k-1} = \mathbf{x}^j) \\ &= P\left(\mathbf{n}_k = \mathbf{x}^i \ominus \mathbf{f}(\mathbf{x}^j)\right),\end{aligned} \quad (12)$$

for $i, j = 1, \ldots, 2^d$. The prediction matrix is entirely a function of the state noise distribution. Additionally, given a value of the observation vector $\mathbf{y}$, let the *update matrix* $T_k(\mathbf{y})$, also of size $2^d \times 2^d$, be a diagonal matrix defined by the observation model (4):

$$\begin{aligned}(T_k(\mathbf{y}))_{jj} &= p_{\mathbf{Y}_k}\left(\mathbf{y} \mid \mathbf{X}_k = \mathbf{x}^j\right) \\ &= p_{\mathbf{v}_k}\left(\mathbf{y} \diamondminus \mathbf{h}(\mathbf{x}^j)\right),\end{aligned} \quad (13)$$

for $j = 1, \ldots, 2^d$, where "$\diamondminus$" is the inverse operation to $\diamondplus$. The update matrix is entirely a function of the observation noise process distribution. If the observation space O is finite, then probability densities become probability mass functions: $p_{\mathbf{Y}_k}(\mathbf{y} \mid \mathbf{X}_k = \mathbf{x}^j) = P(\mathbf{Y}_k = \mathbf{y} \mid \mathbf{X}_k = \mathbf{x}^j)$ an $p_{\mathbf{v}_k}(\mathbf{y} \diamondminus \mathbf{h}(\mathbf{x}^j)) = P(\mathbf{v}_k = \mathbf{y} \diamondminus \mathbf{h}(\mathbf{x}^j))$. In addition, in this case there is a finite number $|O|$ of update matrices, one for each possible value of the observation vector, and these could be computed and stored offline. Finally, define the matrix A of size $d \times 2^d$ as

$$A = \left[\mathbf{x}^1 \cdots \mathbf{x}^{2^d}\right]. \quad (14)$$

The following result gives the algorithm to compute the state MMSE estimator.

Theorem 2. (Boolean Kalman Filter.) *The optimal MMSE estimator of the state process in Theorem 1 can be computed exactly, in recursive fashion, as follows.*

1) *Initialization Step: The initial PDV is given by*

$$\mathbf{\Pi}_{0|0}(i) = P\left(\mathbf{X}_0 = \mathbf{x}^i\right), \quad (15)$$

for $i = 1, \ldots, 2^d$.

For $k \geq 1 = 1, 2, \ldots$, do:

2) *Prediction Step: Given the previous PDV* $\mathbf{\Pi}_{k-1|k-1}$*, the* predicted *PDV* $\mathbf{\Pi}_{k|k-1}$ *is given by*

$$\mathbf{\Pi}_{k|k-1} = M_k \, \mathbf{\Pi}_{k-1|k-1} \,. \tag{16}$$

3) *Update Step: Given the current observation* $\mathbf{Y}_k = \mathbf{y}_k$*, let*

$$\boldsymbol{\beta}_k = T_k(\mathbf{y}_k) \, \mathbf{\Pi}_{k|k-1} \,. \tag{17}$$

The updated *PDV* $\mathbf{\Pi}_{k|k}$ *is obtained by normalizing* $\boldsymbol{\beta}_k$*:*

$$\mathbf{\Pi}_{k|k} = \frac{\boldsymbol{\beta}_k}{||\boldsymbol{\beta}_k||_1} \tag{18}$$

4) *MMSE Estimator Computation Step:*

$$\hat{\mathbf{X}}_k = \overline{E}\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right] = \overline{A\mathbf{\Pi}_{k|k}} \tag{19}$$

with optimal conditional MSE

$$\mathrm{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) = ||\min\{A\mathbf{\Pi}_{k|k}, (A\mathbf{\Pi}_{k|k})^c\}||_1 \,. \tag{20}$$

The previous algorithm, referred to as the *Boolean Kalman Filter*, carries out, in an efficient fashion, the exact propagation of the posterior distribution vectors needed to compute both the MMSE estimator and its optimal MMSE at each time point. The prediction matrix, and the update matrices if O is finite, can be computed offline and stored, which makes the computation very fast.

The normalization step in (18) is not really necessary in order to compute the MMSE estimator and its optimal MMSE at time k. Given the observation history $\mathbf{Y}_1 = \mathbf{y}_1, \ldots, \mathbf{Y}_k = \mathbf{y}_k$, and the initial distribution vector $\mathbf{\Pi}_{0|0}$, let

$$\boldsymbol{\alpha}_k = A\, T_k(\mathbf{y}_k) M_k \cdots T_1(\mathbf{y}_1) M_1 \mathbf{\Pi}_{0|0} \,. \tag{21}$$

Then it is easy to show that

$$E\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right] = \frac{\boldsymbol{\alpha}_k}{||\boldsymbol{\alpha}_k||_1} \,. \tag{22}$$

The lack of normalization can however lead to numerical issues (overflow or underflow) for very long sequences.

V. Inference of Probabilistic Boolean Networks

The Probabilistic Boolean Network (PBN) model is a popular model for gene regulation, where different "contexts" may be active at a given time [4, 5]. The different contexts are determined by external cell stimuli, genetic mutations, and other unobserved processes in the living tissue, which alter the way gene regulation is processed. Gene regulation under each different context is modeled by a different Boolean Network (possibly with perturbation).

The problem of context estimation from sample time series has been so far addressed in the literature by ad-hoc methods [8]. We demonstrate below the application of the BKF in the *optimal* estimation of the context of a small, synthetic PBN.

The PBN consists of two BNs on two variables. The evolution equation (1) for each BN is given by

$$\begin{aligned} \text{BN 1:}\quad & \mathbf{X}_k(1) = \mathbf{X}_{k-1}(1) \text{ NAND } \mathbf{X}_{k-1}(2) \\ & \mathbf{X}_k(2) = \mathbf{X}_{k-1}(1) \text{ XOR } \mathbf{X}_{k-1}(2) \\ \text{BN 2:}\quad & \mathbf{X}_k(1) = \mathbf{X}_{k-1}(1) \text{ AND } \mathbf{X}_{k-1}(2) \\ & \mathbf{X}_k(2) = \mathbf{X}_{k-1}(1) \text{ OR } \mathbf{X}_{k-1}(2) \end{aligned} \tag{23}$$

These BNs have the attractor structure given at the bottom of Figure 2. We can see that BN 1 contains a single cyclic attractor, while BN2 contains three fixed-point attractors. One can expect therefore much faster state transitions under BN 1 than under BN 2.

The unobserved context that selects between the BNs is coded by a third (unobserved) state variable W_k, which determines the context, say $W_k = 0$ for BN 1, and $W_k = 1$ for BN 2. We will assume for this example that the context selector evolves independently of the other state variables, but this assumption can be easily removed. The state model (2) that describes the PBN can then be written as:

$$\begin{bmatrix} \mathbf{X}_k \\ W_k \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{X}_{k-1}; W_k) \\ W_{k-1} \end{bmatrix} \oplus \begin{bmatrix} \mathbf{n}_k \\ r_k \end{bmatrix}, \tag{24}$$

where $\mathbf{f}(\cdot; W_k = 0)$ is the network function for BN 1 and $\mathbf{f}(\cdot; W_k = 1)$ is the network function for BN 2, while the state noise processes are assumed to have independent components, with $\mathbf{n}_k(1), \mathbf{n}_k(2) \sim \text{Bernoulli}(p)$ and $r_k \sim \text{Bernoulli}(c)$, such that $p = 0.05$ and $c = 0.02$. The parameters p and c give the intensity of the noise. Notice that one can assume $p < \frac{1}{2}$, since otherwise the situation is equivalent to having the negated logic functions, with noise intensity $1 - p$. The smaller p is, the more the state transition is governed by the BNs, while $p = \frac{1}{2}$ corresponds to a random state vector. The parameter c is the *context switching probability* of the PBN and determines the chance that the context will switch from BN 1 to BN 2 at a given moment of time. Notice that the context process is not independent; i.e., W_k is not independent of W_l for $k \neq l$. The state model for the PBN is represented in the diagram at the top of Figure 2.

Finally, the observations in this example are assumed to be binary, with observation model (4) given by

$$\mathbf{Y}_k = \begin{bmatrix} \mathbf{X}_k(1) \\ \mathbf{X}_k(2) \end{bmatrix} \oplus \mathbf{v}_k \,, \tag{25}$$

so that only the state variables $\mathbf{X}_k(1)$ and $\mathbf{X}_k(2)$ are directly observable, under noisy conditions. The observation noise process is assumed to have independent components, with $\mathbf{v}_k(1), \mathbf{v}_k(2) \sim \text{Bernoulli}(q)$, where $q = 0.05$.

Figure 3 display realizations of the observation process variables $\mathbf{Y}_k(1)$ and $\mathbf{Y}_k(2)$, the context selector process W_k, the optimal MMSE estimator $\hat{W}_k$ computed by application of the BKF to the observation process, and the conditional expectation $\eta_k = E\left[W_k | \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right]$, respectively. Note that the observations present fast transitions whenever the (unobserved) context of the system is BN 1, whereas transitions are slower when the context is BN 2, in accordance with the previous discussion on the distinct attractor structure of the BNs. This could form the basis for an ad-hoc context inference method [8]; instead, the BKF provides the *optimal* inference procedure, according to MSE. We can observe in Figure 3 that $\hat{W}_k$ is able to track the true signal W_k well. It follows from Theorem 2 that $\hat{W}_k = \overline{\eta_k}$, i.e., $\hat{W}_k$ is obtained by thresholding η_k around $\eta_k = \frac{1}{2}$, with optimal estimation MSE for $\hat{W}_k$ given by $\frac{1}{2} - |\eta_k - \frac{1}{2}|$. The MSE is maximal when η_k

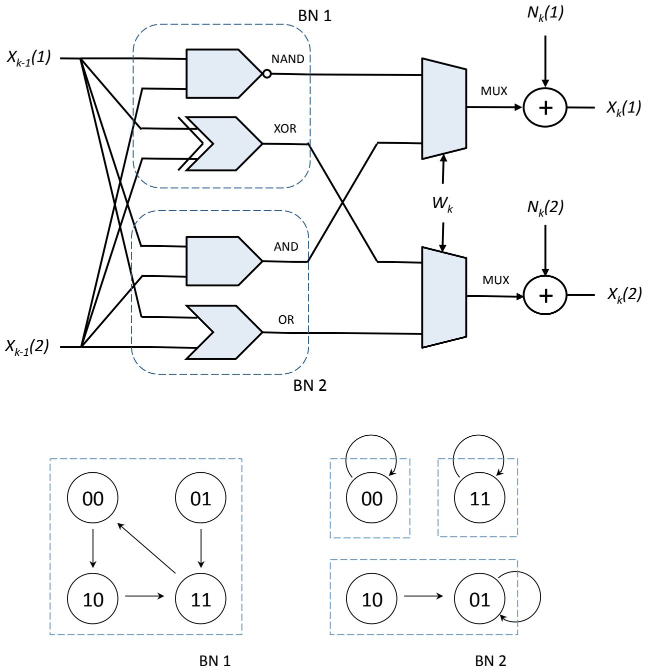


Fig. 2. Example of PBN with two component Boolean Networks (BNs) Top: Block diagram corresponding to state model in (24). The context process W_k determines which one of the two BNs are active at time k; this corresponds to the action of a 2-1 multiplexer (MUX) with selector input W_k. Bottom: State-space diagrams for the two component BNs. Dashed-line boxes indicate the basis of attraction containing each attractor. One can see that BN1 contains a cyclic attractor, which introduces fast state transitions.

is close to $\frac{1}{2}$, and minimal when η_k is close to either 0 or 1. We can therefore interpret $|\eta_k - \frac{1}{2}|$ as the *confidence* in the prediction of the context. It can be seen that the confidence becomes small near a transition of the context. In addition, there can be a lag before $\hat{W}_k$ responds to a transition in W_k (e.g., the transition at $k = 160$). During this lag the BKF must "accumulate" enough contradictory observations in order to change the internal estimate of the context; this is reminiscent of a small Kalman gain in the linear KF case, when the filter "distrusts" the observations. At other times, $\hat{W}_k$ can respond to transitions in W_k quickly (e.g., $k = 275$).

VI. Conclusions

This paper introduced a novel state-space signal model for discrete-time Boolean dynamical systems. The proposed model includes as special cases several distinct stochastic models for Boolean dynamical systems currently in use.

Further extensions and issues remain to be addressed. For example, for large networks with many variables and complex network functions, the exact computation of the BKF may become prohibitive, necessitating the use of approximate, Monte-Carlo sequential methods, such as the Particle Filter [9].

Another interesting question is to study the steady-state behavior of the BKF, and how it relates to the attractor structure of the BNs and the steady-state distribution of the state vector process, since the latter is an ergodic Markov Chain (given minimal requirements on the state noise process).

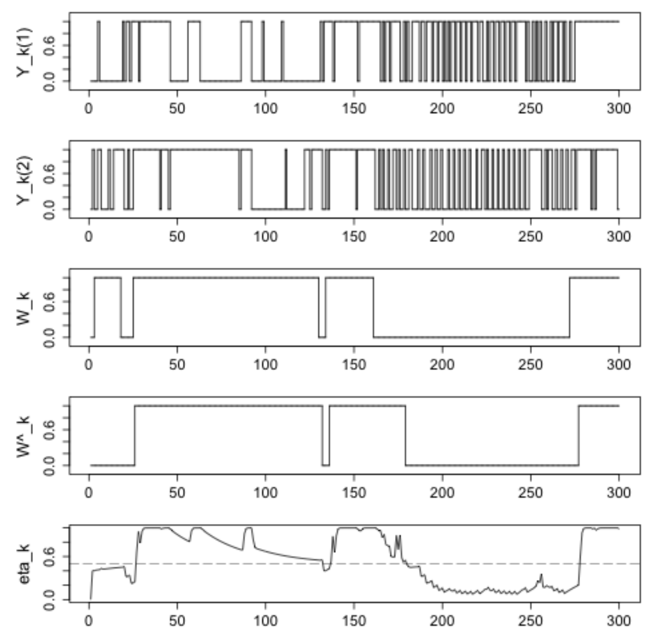


Fig. 3. Sample time series for PBN example.

A very important topic is the issue of incomplete information about the network functions, and how to recover the missing information using the observations, i.e., the problems of system identification and parameter estimation.

Finally, external inputs can be easily included in the state equation (2) by modifying it to

$$\mathbf{X}_k = \mathbf{f}(\mathbf{X}_{k-1}) \oplus \mathbf{u}_k \oplus \mathbf{n}_k \,, \tag{26}$$

where $\mathbf{u}_k$ is a deterministic forcing term. The problem is how to design optimally the input so as to drive the evolution of the system towards desired states. This topic has been studied for the PBN model [4].

Future work will address these and other questions related to the application of the BKF.

References

[1] S. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.
[2] N. Wiener, *Cybernetics*. Cambridge, MA: MIT Press, 1948.
[3] Y. Bengio, "Markovian models for sequential data," *Neural Computing Surveys*, vol. 2, pp. 129–162, 1999.
[4] I. Shmulevich and E. Dougherty, *Probabilistic Boolean Networks*. Philadelphia, PA: SIAM, 2010.
[5] I. Schmulevich, E. 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.
[6] U. Braga-Neto and J. E.T.A. Marques, "From functional genomics to functional immunomics: New challenges, old problems, big rewards," *PLoS Computational Biology*, vol. 2, no. 7, p. e81, 2006.
[7] M. B. et al., "Quantitative mass spectrometry in proteomics: a critical review," *Anal Bioanal Chem*, vol. 389, pp. 1017–1031, 2007.
[8] S. Marshall, L. Yu, Y. Xiao, and E. Dougherty, "Inference of a probabilistic boolean network from a single observed temporal sequence," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2007, 2007.
[9] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. New York, NY: Wiley-Interscience, 2006.