Adaptive Particle Filtering for Fault Detection in Partially-Observed Boolean Dynamical Systems

Arghavan Bahadorinejad, *Student Member, IEEE,* Mahdi Imani, *Student Member, IEEE,* and Ulisses Braga-Neto, *Senior Member, IEEE*

***Abstract*—We propose a novel methodology for fault detection and diagnosis in partially-observed Boolean dynamical systems (POBDS). These are stochastic, highly nonlinear, and derivativeless systems, rendering difficult the application of classical fault detection and diagnosis methods. The methodology comprises two main approaches. The first addresses the case when the normal mode of operation is known but not the fault modes. It applies an innovations filter (IF) to detect deviations from the nominal normal mode of operation. The second approach is applicable when the set of possible fault models is finite and known, in which case we employ a multiple model adaptive estimation (MMAE) approach based on a likelihood-ratio (LR) statistic. Unknown system parameters are estimated by an adaptive expectation-maximization (EM) algorithm. Particle filtering techniques are used to reduce the computational complexity in the case of systems with large state-spaces. The efficacy of the proposed methodology is demontrated by numerical experiments with a large gene regulatory network (GRN) with stuck-at faults observed through a single noisy time series of RNA-seq gene expression measurements.**

***Index Terms*—Fault Detection and Diagnosis, Partially-Observed Boolean Dynamical System, Particle Filtering, Expectation-Maximization, Multiple Model Adaptive Estimation, Gene Regulatory Network, RNA-Seq Gene Expression.**

I. Introduction

FAULT detection and diagnosis in gene regulatory networks (GRN) is a problem of current interest in systems biology [1]–[4], since the molecular basis of many diseases, particularly cancer, resides in the sudden loss of regulatory power in gene networks due to DNA mutations [5]. The dynamical behavior of GRNs can be effectively modeled by Boolean dynamical systems, also known as Boolean networks, where each gene may be thought to be either in an activated or suppressed transcriptional state [6]–[13]. Indeed, prior biological knowledge often consists of gene activation/inactivation pathway diagrams [14]–[18]. This makes Boolean dynamical systems a very convenient model for GRNs.

However, uncertainty in state transition and the use of modern gene-expression technologies require stochastic Boolean state processes and models for noisy indirect measurement data, respectively. The Partially-Observed Boolean Dynamical Systems (POBDS) [10], [13], [19], [20] model addresses both of these requirements; POBDS are stochastic Boolean dynamical systems indirectly observed through a single time series of noisy measurements. Several tools for the POBDS model have been developed in recent years. The optimal filter and smoother based on MMSE criterion, called the Boolean Kalman Filter (BKF) and Boolean Kalman Smoother (BKS), were introduced in [10] and [19], respectively. In addition, schemes for simultaneous state and parameter estimation and their particle filter implementations were developed in [13], [21]. Other tools include sensor selection [22], fault detection [20], and control [23]–[25]. Most of these tools are freely available through an open-source R package called "BoolFilter" [26].

The problem of fault detection and diagnosis (also known as "fault detection and identification") has been studied extensively in many diverse areas of Engineering [27]–[36]. However, classical fault detection and diagnosis methods rely on either system linearity or linearizability assumptions. However, Boolean dynamical systems are highly nonlinear and derivativeless, and thus not easily linearizable. This renders difficult the application of classical fault detection and identification methods. To address this problem, we propose two approaches for fault detection and diagnosis using the POBDS model: the first is meant for the case when only the normal mode of operation is known, and is based on an innovations filter (IF), while the second applies to the case when the set of possible fault models is finite and known, in which case we employ a nonlinear version of the multiple model adaptive estimation (MMAE) approach [37] in combination with a likelihood-ratio (LR) statistic.

The problem of fault detection in Boolean dynamical systems has been addressed previously [1], [20]. In contrast to [1], we do not assume a deterministic system, nor do we assume that the Boolean gene states are directly observable. Instead, we employ the POBDS model to address the more realistic case of an observational layer consisting of a single time series of noisy measurements of the gene states. The methodology in [20] was also developed for POBDS models, but it suffers from two main limitations: (1) it becomes too computationally intensive in the case of large GRNs and (2) it assumes that all system parameters, including the network topology, are known. In this paper, we extend the approach in [20], addressing each of these difficulties. To address the computational difficulty with large systems, we employ an approximate sequential Monte-Carlo (SMC) algorithm for optimal filtering, called the Auxiliary Particle Filtering-Boolean Kalman Filter (APF-BKF) [21], which is adapted here for fault diagnosis and detection purposes. To estimate unknown

M. Imani and U. Braga-Neto are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, 77843. E-mails: mahdy.imani@gmail.com and ulisses@tamu.edu

A. Bahadorinejad is with Apple Inc., Cupertino, CA, 95014. E-mail: arghavan@apple.com

system parameters, we assume an initial (short) fault-free time interval at the start of system operation, and apply an adaptive particle filtering technique based on the expectation-maximization (EM) algorithm for maximum-likelihood system identification [21]. As mentioned previously, these tools are used to obtain an innovations filter in case of unknown faults, and a MMAE approach based on a likelihood-ratio statistic for the case when the set of fault models is finite and known. The efficacy of the proposed methodology is demonstrated via numerical experiments using a Boolean model of a large cell cycle GRN with stuck-at faults that model molecular events commonly found in cancer.

II. Methods

A. Partially-Observed Boolean Dynamical Systems

In this section, the POBDS model is briefly introduced. It consists of a state model that describes the evolution of the Boolean dynamical system, which includes the system input, and an observation model that relates the state to the system output (measurements). More details can be found in [10], [13].

1) State Model: Assume that the system is described by a *state process* $\{\mathbf{X}_k; k = 0, 1, \ldots\}$, where $\mathbf{X}_k \in \{0,1\}^d$ represents the activation/inactivation state of the genes at time k. The states are assumed to be updated at each discrete time through the following nonlinear signal model:

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

for $k = 1, 2, \ldots$ where $\mathbf{u}_k \in \{0,1\}^d$ is the input at time k, $\mathbf{f} : \{0,1\}^d \times \{0,1\}^d \to \{0,1\}^d$ is a Boolean function called the *network function*, "$\oplus$" indicates componentwise modulo-2 addition, and $\mathbf{n}_k \in \{0,1\}^d$ is Boolean transition noise. The noise process $\{\mathbf{n}_k; k = 1, 2, \ldots\}$ is assumed to be "white" in the sense that the noises at distinct time points are independent random variables. We also assume that noise processes are independent of each other and independent of the initial state $\mathbf{X}_0$. We assume that the components of $\mathbf{n}_k$ are i.i.d., with $P(\mathbf{n}_k(i) = 1) = p$, for $i = 1, \ldots, d$. (The general non-i.i.d. case can be similarly handled, at the expense of introducing more parameters.) The parameter $0 \le p \le 1/2$ corresponds to the amount of "perturbation" to the Boolean state process; the case $p = 1/2$ corresponds to maximum uncertainty.

The model adopted here for the network function $\mathbf{f}$ in equation (1) is motivated by gene pathway diagrams commonly encountered in biomedical research [13]:

$$f(\mathbf{x}, \mathbf{u}) = \begin{cases} 1 & \text{if } \sum_{j=1}^{d} a_{ij}\mathbf{x}(j) + b_i + \mathbf{u}(i) > 0, \\ 0 & \text{if } \sum_{j=1}^{d} a_{ij}\mathbf{x}(j) + b_i + \mathbf{u}(i) \le 0, \end{cases} \tag{2}$$

where $a_{ij} = 1$ if gene j activates gene i, $a_{ij} = -1$ if gene j inhibits gene i, $a_{ij} = 0$ if there is no regulation from gene j to gene i, and parameter b_i specifies the regulation bias. From (2), a target gene is activated if the number of activation inputs exceed the number of inhibition inputs, otherwise it is inactivated. For target gene i, and assuming that $\mathbf{u}(i) = 0$, for simplicity, we can see that tie-breaking is determined by the bias parameter b_i. If this parameter has a small positive value (say, $b_i = +1/2$), then a tie is decided in favor of activation. If it is zero, then a tie is decided in favor of inactivation.

2) Measurement Model: The second component of the signal model is the measurement model. In most real-world applications, the system state is only partially observable, and distortion is introduced in the observations by sensor noise — this is certainly the case with gene expression data.

Let $\mathbf{Y}_k = (\mathbf{Y}_k(1), \ldots, \mathbf{Y}_k(d))$ be a vector containing the gene expression data at time k, for $k = 1, 2, \ldots$ We consider here the case where gene expression data are obtained by the currently popular *RNA-seq* platform [38]. In this case, $\mathbf{Y}_k(i)$ is the read count corresponding to transcript i, for $i = 1, \ldots, d$. There are multiple methods for modeling RNA-seq reads. Because of the discrete nature of reads, most methods are based on either the negative binomial [13], [39]–[41] or the Poisson distribution [42]. In this study, we choose to use the negative binomial model for the number of reads for each transcript. The negative binomial model has the ability to address overdispersion in the count distributions. We also assume conditional independence of the transcript counts given the state. Under these assumptions,

$$\begin{aligned} P(\mathbf{Y}_k \mid \mathbf{X}_k = \mathbf{x}) &= \prod_{j=1}^{d} P(\mathbf{Y}_k(j) = \mathbf{y}(j) \mid \mathbf{X}_k(j) = \mathbf{x}(j)), \\ &= \prod_{j=1}^{d} \frac{\Gamma(\mathbf{y}(j) + \phi_j)}{\mathbf{y}(j)!\,\Gamma(\phi_j)} \left(\frac{\lambda_j}{\lambda_j + \phi_j}\right)^{\mathbf{y}(j)} \left(\frac{\phi_j}{\lambda_j + \phi_j}\right)^{\phi_j}, \end{aligned} \tag{3}$$

where ϕ_j, λ_j are inverse dispersion parameters that model observation noise and mean read count of transcript j, respectively. Recall that, according to the Boolean state model, there are two possible states for the abundance of transcript j at time k: high ($\mathbf{X}_k(j) = 1$) and low ($\mathbf{X}_k(j) = 0$). Accordingly, we model the parameter λ_j as follows:

$$\log \lambda_j = \log s + \mu_b + \delta_j \mathbf{X}_k(j). \tag{4}$$

The parameter s is the sequencing depth [43], and is assumed here to be common to all transcripts, since a single lane is being modeled. The sequencing depth s accounts for different total numbers of reads produced in the lane and plays a key role, since it determines the approximate range of read counts that is produced. The parameter $\mu_b \ge 0$ accounts for the baseline level of read counts produced in the single lane in the inactivated transcriptional state, which is assumed to be common to all transcripts. On the other hand, $\delta_j \ge 0$ expresses the effect on the observed RNA-seq read count as gene j goes from the inactivated to the activated state, for $j = 1, \ldots, d$. In summary, the measurement model parameters are the sequencing depth s, the baseline expression level μ_b, the transcript-dependent differential expression levels δ_j, for $j = 1, ..., d$, and the transcript-dependent inverse dispersion parameters ϕ_j, for $j = 1, ..., d$.

B. Auxiliary Particle Filter Implementation of the Boolean Kalman Filter (APF-BKF)

The minimum mean-square error (MMSE) filtering problem consists of, given observations $\mathbf{Y}_{1:k} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$, finding

an estimator $\hat{\mathbf{X}}_k = h(\mathbf{Y}_{1:k})$ of the state $\mathbf{X}_k$ that minimizes the mean-square error (MSE),

$$\mathrm{MSE}(\hat{\mathbf{X}}_k \mid \mathbf{Y}_{1:k}) = E\left[||\hat{\mathbf{X}}_k - \mathbf{X}_k||^2 \mid \mathbf{Y}_{1:k}\right], \quad (5)$$

at each value of $\mathbf{Y}_{1:k}$ (so that it also minimizes the frequentist expected MSE over all possible realizations of $\mathbf{Y}_{1:k}$) up to the current time k, for $k = 1, 2, \ldots$

For a vector $\mathbf{v} \in [0,1]^d$, define the *binarized vector* $\overline{\mathbf{v}}(i) = I_{\mathbf{v}(i)>1/2}$ for $i = 1, \ldots, d$. It has been proved [13] that the optimal MMSE filter $\hat{\mathbf{X}}_k^{\mathrm{MS}}$ is given by

$$\hat{\mathbf{X}}_k^{\mathrm{MS}} = \overline{E\left[\mathbf{X}_k \mid \mathbf{Y}_{1:k}\right]}. \quad (6)$$

The optimal Boolean MMSE estimator is called the Boolean Kalman Filter (BKF) [10], [13]. Exact calculation of the BKF requires the computation and storage of $2^d \times 2^d$ transition and update matrices [10]. When the number d of states is large, exact computation of the BKF becomes intractable due to the matrix sizes. In this case, approximate methods must be used, such as sequential Monte-Carlo methods, also known as particle filter algorithms. In this paper, we employ the APF-BKF algorithm, an Auxiliary Particle Filter (APF) Monte-Carlo implementation of the BKF, first proposed in [21], and briefly reviewed next.

In the APF-BKF algoritm, particles $\{\mathbf{x}_{k-1,i}\}_{i=1}^N$ at time $k-1$, with associated normalized weights $\{W_{k-1,i}\}_{i=1}^N$, represent the state posterior distribution at time step $k-1$ via

$$P(\mathbf{X}_{k-1} = \mathbf{x} \mid \mathbf{Y}_{1:k-1}) = \sum_{i=1}^N W_{k-1,i}\, I_{\mathbf{x}=\mathbf{x}_{k-1,i}}, \quad (7)$$

where $I_{\mathbf{x}=\mathbf{x}_{k-1,i}}$ is an indicator which returns 1 if $\mathbf{x} = \mathbf{x}_{k-1,i}$ and 0 otherwise. Since the intensity of the process noise $\mathbf{n}_k$ satisfies $p < 1/2$, one can write:

$$\mu_{k,i} = \mathrm{Mode}[\mathbf{f}(\mathbf{x}_{k-1,i}) \oplus \mathbf{n}_k] = \mathbf{f}(\mathbf{x}_{k-1,i}), \quad (8)$$

for $i = 1, \ldots, N$. Now, two stage weights should be computed as follows:

- The first-stage weights $\{V_{k,i}\}_{i=1}^N$ are computed as:

$$V_{k,i} = P(\mathbf{Y}_k \mid \mu_{k,i}) W_{k-1,i}, \quad (9)$$

  for $i = 1, \ldots, N$.
- For the second-stage weights, one first needs to perform the following sampling process:

$$\{\zeta_{k,i}\}_{i=1}^N \sim \mathrm{Cat}(\{V_{k,i}\}_{i=1}^N), \quad (10)$$

  and then compute the particles and weights as:

$$\begin{aligned} \mathbf{x}_{k,i} &= \mu_{k,\zeta_{k,i}} \oplus \mathbf{n}_k \sim P(\mathbf{X}_k \mid \mathbf{x}_{k-1,\zeta_{k,i}}), \\ \tilde{W}_{k,i} &= \frac{P(\mathbf{Y}_k \mid \mathbf{x}_{k,i})}{P(\mathbf{Y}_k \mid \mu_{k,\zeta_{k,i}})}. \end{aligned} \quad (11)$$

  for $i = 1, \ldots, N$. Here, "Cat" stands for the categorical (discrete) distribution.

Using the first and second-stage weights, it is shown in [21] that an unbiased estimator of the unnormalized posterior probability of the state given the observations up to time k is given by:

$$||\hat{\boldsymbol{\beta}}_k||_1 = \left(\frac{1}{N}\sum_{i=1}^N V_{k,i}\right)\left(\frac{1}{N}\sum_{i=1}^N \tilde{W}_{k,i}\right). \quad (12)$$

Given the normalized second-stage weights $W_{k,i} = \tilde{W}_{k,i}/\sum_{j=1}^N \tilde{W}_{k,j}$, $i = 1, \ldots, N$, one can write

$$E[\mathbf{X}_k \mid \mathbf{Y}_{1:k}] \approx \mathbf{z}_k = \sum_{i=1}^N W_{k,i}\, \mathbf{x}_{k,i}. \quad (13)$$

From (6), it follows that the MMSE state estimator at time step k can be approximated as $\hat{\mathbf{X}}_k^{\mathrm{MS}} \approx \overline{\mathbf{z}_k}$. The entire procedure is summarized in Algorithm 1. For more information, see [21].

Algorithm 1 APF-BKF: Auxiliary Particle Filter implementation of the Boolean Kalman Filter [21]

```
x_{0,i} ~ Π_{0|0}, W_{0,i} = 1/N, for i = 1, ..., N.
for k = 1, 2, ..., do
    for i = 1 to N do
        μ_{k,i} = f(x_{k-1,i}).
        V_{k,i} = P(Y_k | μ_{k,i}) W_{k-1,i}.
    end for
    {ζ_{k,i}}_{i=1}^N ~ Cat({V_{k,i}}_{i=1}^N).
    for i = 1 to N do
        x_{k,i} = μ_{k,ζ_{k,i}} ⊕ n_k.
        W̃_{k,i} = P(Y_k|x_{k,i}) / P(Y_k|μ_{k,ζ_{k,i}}).
    end for
    ||β̂_k||_1 = (1/N Σ_{i=1}^N V_{k,i}) (1/N Σ_{i=1}^N W̃_{k,i}).
    W_{k,i} = W̃_{k,i} / Σ_{j=1}^N W̃_{k,j}, i = 1, ..., N.
    z_k = Σ_{i=1}^N W_{k,i} x_{k,i}.
    X̂_k^MS = z̄_k.
end for
```

C. Parameter Estimation

Suppose that the POBDS model is incompletely specified; i.e., the transition noise intensity p, or the parameters of the RNA-seq model in (3) and (4) may be unknown or only partially known. Let the missing information be encoded into a finite-dimensional vector $\theta \in \Theta$, where $\Theta \subseteq \mathbb{R}^m$.

In [13], a maximum likelihood adaptive filter was developed for simultaneous state and parameter estimation for POBDS with unknown continuous parameters. The method is based on the well-known expectation maximization (EM) algorithm [44] and the BKF. The EM method attempts to maximize the "complete" log-likelihood function $\log P_\theta(\mathbf{X}_{0:T}, \mathbf{Y}_{1:T})$ in an iterative fashion by estimating a sequence of parameters $\{\theta^{(n)}; n = 0, 1, \ldots\}$. Given the current estimate $\theta^{(n)}$, the

algorithm obtains the next estimate $\theta^{(n+1)}$ in the sequence by computing the Q function (E-step):

$$Q(\theta, \theta^{(n)}) \;=\; \sum_{\mathbf{X}_{0:T}} \log P_\theta(\mathbf{X}_{0:T}, \mathbf{Y}_{1:T})\, P_{\theta^{(n)}}(\mathbf{X}_{0:T} \mid \mathbf{Y}_{1:T}) \tag{14}$$

and then maximizing it (M-step):

$$\theta^{(n+1)} \;=\; \operatorname*{argmax}_\theta Q(\theta, \theta^{(n)})\,. \tag{15}$$

The exact computation of the E-Step requires applying the BKS, which is computationally expensive and uses a large amount of memory for storing values. The APF-CPMLA-BKS algorithm [13] provides an efficient particle filtering implementation of the EM algorithm for POBDS, which is suitable for large systems with tight time constraints.

The APF-CPMLA-BKS algorithm is briefly explained next. It consists of two main steps:

1) Forward Step: the APF-BKF algorithm is employed to obtain particles $\{\mathbf{x}_{0:T,i}\}_{i=1}^N$, and their associated weights $\{W_{0:T,i}\}_{i=1}^N$, from time 0 to T.
2) Backward Step: the backward procedure simulates N trajectories from the joint conditional distribution:

$$\{\tilde{\mathbf{x}}_{0:T,i}\}_{i=1}^N \;\sim\; P(\mathbf{X}_{0:T} \mid \mathbf{Y}_{1:T}).$$

The obtained trajectories can be used for approximating the Q-function in (14) as:

$$\hat{Q}(\theta, \theta^{(n)}) \;=\; \frac{1}{N} \sum_{i=1}^{N} \Bigg[\log P_\theta(\tilde{\mathbf{x}}_{0,i}) + \sum_{s=1}^{T} \log P_\theta(\tilde{\mathbf{x}}_{s,i} \mid \tilde{\mathbf{x}}_{s-1,i}) + \sum_{s=1}^{T} \log\; P_\theta(\mathbf{Y}_s \mid \tilde{\mathbf{x}}_{s,i}) \Bigg]. \tag{16}$$

The procedure is described in Algorithm 2. Here, "$\xleftarrow{\text{Unique}}$" computes the unique particles with their associated weights or indices. For more information, see [21].

III. Fault Detection and Diagnosis

In this section, we present the proposed algorithms for fault detection and diagnosis. First, we assume that an initial sequence of data of fixed length is available for the system under normal operation. This sequence is used for estimation of the unknown parameters of the system using the APF-CPMLA-BKS algorithm.

Regarding the fault detection process, two options are considered in this paper:

1) Innovations-Filter (IF) Method: The fault models are unknown, and an innovations-based change detection algorithm using a particle filter is proposed — in this case, no fault diagnosis is possible;
2) Likelihood-Ratio (LR) Method: The set of possible fault models is finite and known, and fault detection and diagnosis is accomplished simultaneously using a multiple model estimation approach based on a likelihood-ratio statistic.

Algorithm 2 APF-CPMLA-BKS: APF implementation of the continuous-parameter ML Adaptive BKS [21]

Specify $\theta^{(0)}$ (initial guess) and tolerance $\varepsilon > 0$.
$n \leftarrow -1$.
repeat
  $n \leftarrow n + 1$.
  $\{\mathbf{x}_{0:T,i}, W_{0:T,i}\}_{i=1}^N \leftarrow$ Run APF-BKF tuned to $\theta^{(n)}$.
  $\{\mathbf{x}^u_{k,j}, W^u_{k,j}\}_{j=1}^{F_k} \xleftarrow{\text{Unique}} \{\mathbf{x}_{k,i}, W_{k,i}\}_{i=1}^N$, $k = 0, \ldots, T$.
  Sample $\{\eta_T(i)\}_{i=1}^N \sim \text{Cat}\,(\{W^u_{T,j}\}_{j=1}^{F_T})$.
  Set $\tilde{\mathbf{x}}_{T,i} = \mathbf{x}^u_{T,\eta_T(i)}$, for $i = 1, \ldots, N$.
  for $s = T-1$ to 0 **do**
    $\{\tilde{\mathbf{x}}^u_{s+1,j}, \xi^j_{s+1}\}_{j=1}^{S_{s+1}} \xleftarrow{\text{Unique}} \{\tilde{\mathbf{x}}_{s+1,i}\}_{i=1}^N$.
    for $j = 1$ to S_{s+1} **do**
      $D^j_{s,i} \;=\; W^u_{s,i}\, P(\tilde{\mathbf{x}}^u_{s+1,j} \mid \mathbf{x}^u_{s,i})$, $i = 1, \ldots, F_s$.
      $\{\eta_s(t)\}_{t=1}^{|\xi^j_{s+1}|} \sim \text{Cat}\,(\{D^j_{s,i}\}_{i=1}^{F_s})$.
      $\tilde{\mathbf{x}}_{s,\xi^j_{s+1}(t)} \;=\; \mathbf{x}^u_{s,\eta_s(t)}$, for $t = 1, \ldots, |\xi^j_{s+1}|$.
    end for
  end for
  Find $\hat{Q}(\theta, \theta^{(n)})$ using equation (16).
  Find $\theta^{(n+1)} \;=\; \operatorname{argmax}_\theta \hat{Q}\left(\theta, \theta^{(n)}\right)$.
until $|\theta^{(n+1)} - \theta^{(n)}| > \varepsilon$
$\hat{\theta}^{\text{ML}} = \theta^{(n+1)}$.

A. Fault Detection based on Innovations Filter

It is a well-known fact in the theory of linear Kalman filtering that the innovations of the optimal state estimator constitute a "white noise" sequence [45]. This fact also applies to nonlinear, non-Gaussian systems. For completeness, we give the main result below. Given the history of observations $\mathbf{Y}_{1:k} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$ up to the present time k, let $\mathbf{Y}_k^{\text{MS}} = E\,[\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1}]$ be the *MS-predictable* component of $\mathbf{Y}_k$, with $\mathbf{Y}^{\text{MS}}(j) = E[\mathbf{Y}_1(j) \mid \mathbf{Y}_{1:k-1}]$, for $j = 1, \ldots, d$. The *innovation* $\mathbf{V}_k$ at time k is defined as the MS-unpredictable component of $\mathbf{Y}_k$ as:

$$\begin{aligned} \mathbf{V}_k \;&=\; \mathbf{Y}_k - \mathbf{Y}_k^{\text{MS}} \\ &=\; \mathbf{Y}_k - E\,[\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1}]\;, \end{aligned} \tag{17}$$

for $k = 1, 2, \ldots$. It can be shown [20] that the innovations sequence $\{\mathbf{V}_k; k = 0, 1, \ldots\}$ is zero mean and uncorrelated, which means $E[\mathbf{V}_k] = \mathbf{0}_d$ and $E[\mathbf{V}_k \mathbf{V}_l] = \mathbf{0}_{d\times 0}$, for $k = 1, 2 \ldots$, and $k \neq l$.

One can write $\mathbf{Y}_k^{\text{MS}}$ as:

$$\begin{aligned} \mathbf{Y}_k^{\text{MS}} \;&=\; E\,[\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1}] \\ &=\; E\,[E\,[\mathbf{Y}_k \,|\, \mathbf{X}_k, \mathbf{Y}_{1:k-1}] \mid \mathbf{Y}_{1:k-1}] \\ &=\; E[\,E[\mathbf{Y}_k \mid \mathbf{X}_k] \mid\; \mathbf{Y}_{1:k-1}] \\ &=\; \sum_{i=1}^{2^d} E[\mathbf{Y}_k \mid \mathbf{X}_k = \mathbf{x}^i]\, P\left(\mathbf{X}_k = \mathbf{x}^i \mid\; \mathbf{Y}_{1:k-1}\right). \end{aligned} \tag{18}$$

Using (3), $E[\mathbf{Y}_k \mid \mathbf{X}_k = \mathbf{x}^i]$ can be written as:

$$E[\mathbf{Y}_k(j) \mid \mathbf{X}_k = \mathbf{x}^i] \;=\; \lambda_j \;=\; s\, \exp(\mu_b + \delta_j(\mathbf{x}^i(j)))\,, \tag{19}$$

for $j = 1, \ldots, d$ and $i = 1, \ldots, 2^d$. It is clear that the computation of $\mathbf{Y}_k^{\mathrm{MS}}$ requires performing a BKF for posterior computation. However, we are dealing here with large systems, to which the BKF cannot be applied due to the large computational complexity. In the following paragraph, we show how the whiteness of the innovations sequence can be tested approximately using the particles and weights computed by the APF-BKF.

Let $\{W_{k-1,i}, \mathbf{x}_{k-1,i}\}_{i=1}^N$ be the particles and their associated weights at time step $k-1$ computed by the APF-BKF algorithm from the sequence of measurements $\mathbf{Y}_{1:k-1}$. The predicted particles at time step k, before observing measurement $\mathbf{Y}_k$, can be computed as:

$$\mathbf{x}_{k,i}^p \;=\; \mathbf{f}(\mathbf{x}_{k-1,i}, \mathbf{u}_k) \oplus \mathbf{n}_k^i, \tag{20}$$

for $i = 1, \ldots, N$. Using the predicted particles at time step k, and also the particles and their associated weights at time step $k-1$, equation (18) can be approximated as:

$$\begin{aligned}
\mathbf{Y}_k^{\mathrm{MS}} &\approx \sum_{i=1}^{N} E[\mathbf{Y}_k \mid \mathbf{X}_k = \mathbf{x}_{k,i}^p] P\left(\mathbf{X}_k = \mathbf{x}_{k,i}^p \mid \ \mathbf{Y}_{1:k-1}\right) \\
&= \sum_{i=1}^{N} E[\mathbf{Y}_k \mid \mathbf{X}_k = \mathbf{x}_{k,i}^p]\, P\left(\mathbf{X}_k = \mathbf{x}_{k,i}^p \mid \mathbf{X}_{k-1} = \mathbf{x}_{k-1,j}\right) \\
&\qquad \times P\left(\mathbf{X}_{k-1} = \mathbf{x}_{k-1,j} \mid \mathbf{Y}_{1:k-1}\right) \\
&= \sum_{i=1}^{N} s \exp(\mu_b + \delta_j \mathbf{x}_{k,i}^p) \sum_{j=1}^{N} P(\mathbf{x}_{k,i}^p \mid \mathbf{x}_{k-1,j})\, W_{k-1,j}.
\end{aligned} \tag{21}$$

Assuming normal system operation, the APF-BKF is run, and the residue at time k can be computed as:

$$\mathbf{e}_k \;=\; \mathbf{Y}_k - \mathbf{Y}_k^{\mathrm{MS}}, \tag{22}$$

for $k = 1, 2, \ldots$

From the previous results, the sequence $\{\mathbf{e}_k; k = 1, 2, ...\}$ is the innovations sequence, and therefore "white noise", provided that the normal-operation model, assumed by the APF-BKF, matches the actual model producing the data. Therefore, a fault detection method results from testing the hypothesis that the residue sequence is white. In this paper, similarly to [20], a chi-square test is applied to L lags of the samples auto-correlation function based on an "window" of L_D observations preceding the current time k_0. To improve false-positive error rates, we adopt the fault certification step described in [46], in which a fault is signaled only if the number of false detections over a window of L_C observations preceding k_0 exceeds a threshold t_C.

B. Fault Detection and Diagnosis based on Likelihood Ratios

If the set of possible faults is finite and known beforehand, we propose an approach for simultaneous fault detection and diagnosis, which consists of an online likelihood-ratio cumulative sum (CUSUM) procedure [47] within a multiple model adaptive estimation (MMAE) scheme, where a bank of particle filters runs in parallel, each tuned to a different fault model [21].

Let $\{\theta^1, \ldots, \theta^M\}$ parametrize the set of all possible fault models, and let θ^0 correspond to the normal-operation model. Given a sequence of measurements $\mathbf{Y}_{1:k}$, the cumulative sum of the log-likelihood ratio can be computed recursively for each fault model as follows:

$$\begin{aligned}
S_k^{\theta^m} &= \log \frac{P_{\theta^m}(\mathbf{Y}_{1:k})}{P_{\theta^0}(\mathbf{Y}_{1:k})} \\
&= \log \frac{P_{\theta^m}(\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1}) P_{\theta^m}(\mathbf{Y}_{1:k-1})}{P_{\theta^0}(\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1}) P_{\theta^0}(\mathbf{Y}_{1:k-1})} \\
&= S_{k-1}^{\theta^m} + \log \frac{P_{\theta^m}(\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1})}{P_{\theta^0}(\mathbf{Y}_k \mid \mathbf{Y}_{1:k-1})} \\
&= S_{k-1}^{\theta^m} + \log \frac{||\hat{\boldsymbol{\beta}}_k^{\theta^m}||_1}{||\hat{\boldsymbol{\beta}}_k^{\theta^0}||_1},
\end{aligned} \tag{23}$$

for $m = 1, \ldots, M$, where $\hat{\boldsymbol{\beta}}_k^{\theta^m}$ is the unnormalized posterior probability of the APF-BKF tuned to parameter θ^m at time step k. The likelihood-ratios are thus computed by a bank of $M+1$ APF-BKFs running in parallel, as in the MMAE algorithm [21].

The typical behavior of the log-likelihood ratio $S_k^{\theta^m}$ for a faulty system is a negative drift before change, and a positive drift after change [47]. Therefore, we would like to track the difference between the value of the log-likelihood ratio and its current minimum value:

$$g_k^{\theta^m} \;=\; S_k^{\theta^m} - \min_{1 \le l \le k} S_l^{\theta^m}, \tag{24}$$

for $m = 1, \ldots, M$. A fault is detected at the first time step at which the value of the adaptive cumulative log-likelihood ratio $g_k^{\theta^m}$ exceeds a a conveniently selected threshold h for one of the fault models θ^*, at which time the fault is also diagnosed as θ^* (thus, fault detection and diagnosis are processed simultaneously). Formally, the detection time k_D is given by:

$$k_d \;=\; \min\{k \mid g_k^{\theta^m} > h \text{ for some } m = 1, \ldots, M\} \tag{25}$$

and the fault is diagnosed as

$$\hat{\theta}^{\mathrm{LR}} \;=\; \underset{m=1,\ldots,M}{\operatorname{argmax}}\; g_{k_d}^{\theta^m}. \tag{26}$$

Notice that this has the advantage of breaking ties in case two or more models signal a fault at time k_d.

The likelihood-ratio procedure is described in Algorithm 3.

IV. Results and Discussion

Several metrics can be defined to assess the performance of the proposed fault detection and diagnosis algorithms; e.g. see [46]. The error rates and average lag times are used here for our analysis. Let K be the total length of the time series of observations, $k_0 < K$ be the time at which a fault occurs, and k_d be the time at which a fault is first detected or diagnosed. If $k_d < k_0$, a *false alarm* has occurred. If $k_d \ge k_0$, a fault has been correctly detected, and we are interested in how fast the fault is detected, i.e., how small $k_d - k_0$ is. The average of these quantities has been used in our numerical experiments.

In our simulation, we generate T time series of length K. Without loss of generality, the true time k_0 at which the fault

Algorithm 3 Likelihood-ratio procedure for fault detection and diagnosis based on a bank of APF-BKFs

$S_0^{\theta^m} = 0$, for $m = 1, \ldots, M$
for $k = 1, 2, \ldots,$ **do**
  Compute $\hat{\boldsymbol{\beta}}_k^{\theta^0}$ using an APF-BKF tuned to θ^0.
  for $m = 1, \ldots, M$ **do**
    Compute $\hat{\boldsymbol{\beta}}_k^{\theta^m}$ using an APF-BKF tuned to θ^m.
    $S_k^{\theta^m} = S_{k-1}^{\theta^m} + \log \frac{\|\hat{\boldsymbol{\beta}}_k^{\theta^m}\|_1}{\|\hat{\boldsymbol{\beta}}_k^{\theta^0}\|_1}$.
    $g_k^{\theta^m} = S_k^{\theta^m} - \min_{1 \le l \le k} S_l^{\theta^m}$,
  end for
  if $g_k^{\theta^m} > h$ for some $m = 1, \ldots, M$ **then**
    $k_D = k$
    $\hat{\theta}^{\mathrm{LR}} = \mathrm{argmax}_{m=1,\ldots,M}\, g_{k_D}^{\theta^m}$.
    Stop.
  end if
end for

occurs is assumed to be fixed for all time series. The *false detection rate* (FDR) is defined as:

$$\mathrm{FDR} = \frac{1}{T} \sum_{t=1}^{T} I_{k_d^t < k_0}\,, \tag{27}$$

where k_d^t is the time when a fault is detected for time series t, for $t = 1, \ldots, T$, and I_A returns 1 if condition A is true and 0, otherwise. The *average time until correct detection* (ATCD) is defined as:

$$\mathrm{ATCD} = \frac{1}{T_d} \sum_{t=1}^{T} (k_d^t - k_0) I_{k_d^t \ge k_0}\,, \tag{28}$$

where $T_d = \sum_{t=1}^{T} I_{k_d^t \ge k_0}$ is the number of time series for which a correct detection has been made. Finally, let θ^* and $\hat{\theta}_t^{\mathrm{LR}}$ be the true model and diagnosed fault model for the tth trajectory, respectively. The *false diagnosis rate* (FDGR) is defined as:

$$\mathrm{FDGR} = \frac{1}{T_d} \sum_{t=1}^{T} I_{\hat{\theta}_t^{\mathrm{LR}} \neq \theta^*} I_{k_d^t \ge k_0}\,. \tag{29}$$

We employ in our experiments the GRN in [48], which models the dynamical response of the budding yeast cell cycle to osmotic and pheromone stress. The network contains 14 genes and two inputs, Far1 and Hog. The regulatory relationships between genes in this network are shown in Figure 1. We use these relationships to set the parameters in the network model in (2). All biases b_i are set to zero.

All numerical experiments use the parameter values displayed in Table I. The transition noise parameter p, the baseline expression μ_b, and differential expression δ_i, $i = 1, 2, \ldots, 14$, are assumed to be unknown and are estimated

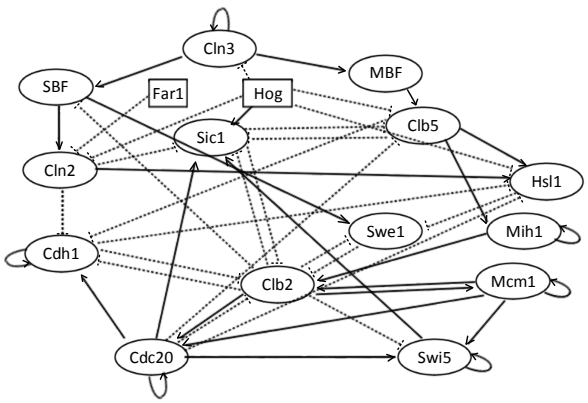

Fig. 1. Boolean model of the regulatory network of cell cycle under two different extracellular signals osmatic stress and the α-factor.

using an initial fault-free sequence of length K_0, as described previously (initially, $K_0 = 50$ is fixed, but later it is varied to examine its effect on the proposed fault detection and diagnosis methods). The true fault time k_0 is set to $K_0 + 200$ for fault detection based on the innovations filter (IF), when there are no knowledge of the fault models (and thus no fault diagnosis), and to $K_0 + 50$ for fault detection and diagnosis based on likelihood-ratios (LR), when the fault models are assumed to be known. The length of the time series is set to $K = 400$ for the IF method and $K = 100$ for the LR method. The reason for these choices is the fact that the IF method assumes less information (no knowledge about the possible fault models) and therefore requires more data. The sequencing depth s (which is instrument-dependent and usually available) is assumed to be known, and so are the inverse dispersion parameters $\phi_i = 2$, $i = 1, 2, \ldots, 14$.

Table II displays the results with $K_0 = 50$. One can see that the false detection rate (FDR) and average time until correct detection (ATCD) are generally larger for the IF method than for the LR method. The reason of these results can again be justified by the availability of prior information on the possible fault models in the LR method. In addition, the LR method requires a smaller window for detection in comparison with the IF method. One can observe that in nearly all cases the results improve substantially as the number of NGS reads increases, as expected. Some faults can be harder to detect than others. For example, the Swi5 gene under a stuck-at-1 fault, and SBF gene under stuck-at-0 present an elevated FDR. This is a consequence of the specific structure of this Boolean network.

Figure 2 displays the adaptive cumulative log-likelihood ratio g_k in (24) for the LR method as a function of time (excluding the initial K_0 time points used for system identification). The true fault occurs at time point 50 in all plots. It can be seen that in nearly all cases, the adaptive cumulative log-likelihood ratio starts to increase at the time the fault occurs.

Table III displays parameter estimation results obtained by the APF-CPMLA-BKS algorithm using the fault-free sequence of length $K_0 = 50$. As expected, the results converge to the true parameter values for large sequencing depths.

Finally, the effect of the length K_0 of the initial fault-free

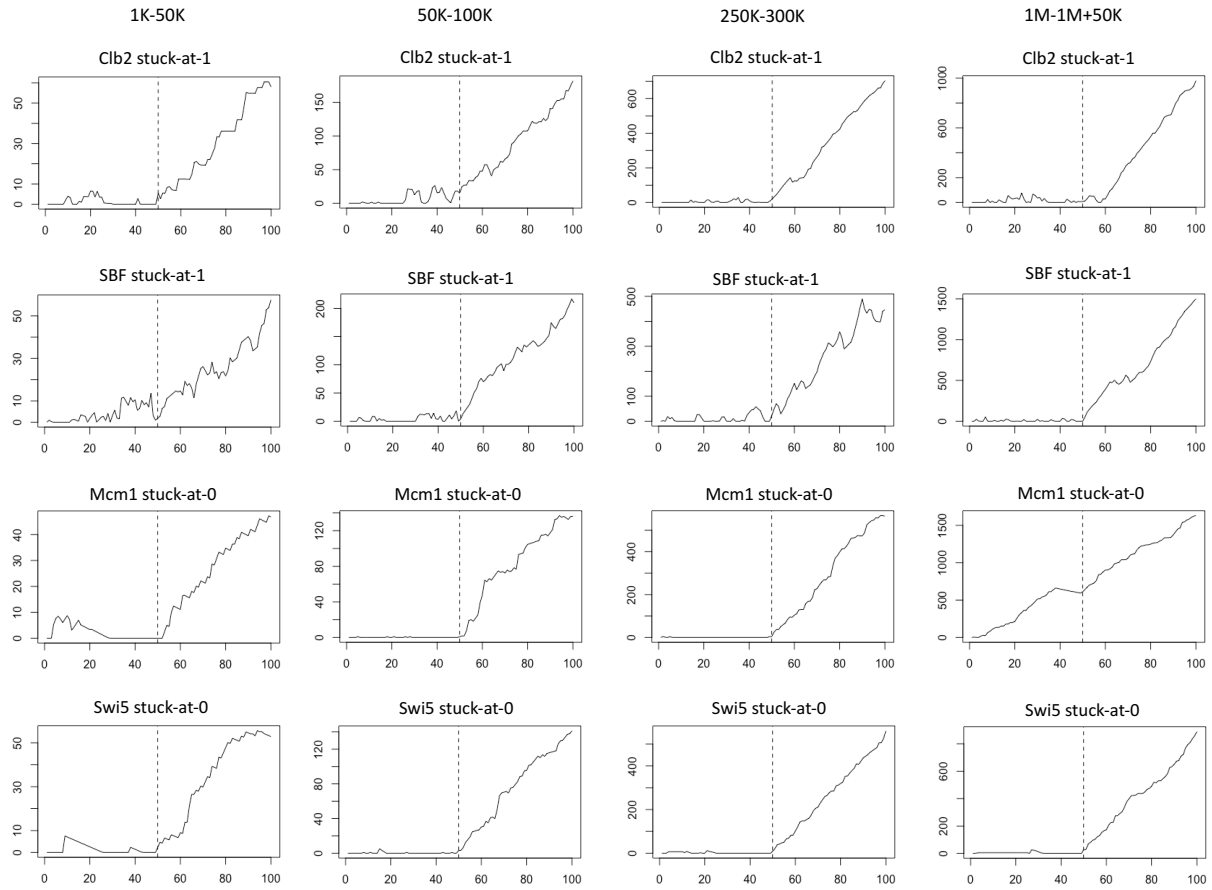

Fig. 2. Cumulative log-likelihood ratios for the LR method as a function of time for Hog=1 and Far1=0. The true fault occurs at time 50 in all plots.

sequence for system identification on the performance of the proposed LR method was examined. The time at which the fault occurs is K_0+50. The false diagnosis rate (FDGR) of the proposed LR method for various values of K_0 is displayed in Figure 3. It can be seen that as K_0 increases, the performance of the LR fault detection and diagnosis method improves, as expected, since parameter estimation becomes more accurate. There is a point of diminishing returns, which varies according to the fault, beyond which the FDGR rates converge to fixed values, and additional data for system identification does not increase performance significantly. Notice also that the limiting value for the FDGR is not always zero.

V. Conclusion

In this paper, fault detection and diagnosis of partially-observed Boolean dynamical systems (POBDS) has been discussed. The parameters of the system are assumed to be partially-known. First, the particle-based maximum likelihood adaptive filter was used for estimating the unknown parameters of the system using an initial fault-free samples. Then, the estimated parameters were used as true unknown parameters of the system for fault detection and diagnosis process. Two cases were considered in this paper: 1) In the first case, the fault models were assumed to be unknown and the whiteness of innovation auxiliary particle filter is used for detecting the fault in the system. 2) In the second case, the fault models were assumed to be known and the maximum likelihood ratio technique was developed for simultaneous detection and diagnosis processes. The efficacy of the proposed methodologies was demonstrated via numerical experiments using synthetic RNA-seq gene expression data from a realistic GRN model of the cell cycle with stuck-at mutational events. We remark that it was not our purpose to develop a finished product for fault detection in GRNs using experimental data. The challenges associated with designing experiments to obtain such data go beyond the scope of the present paper, and they were left for future work. Nevertheless, the results obtained with the synthetic data indicate that the proposed methods are promising in monitoring and diagnosing biological changes at the transcriptomic level.

Acknowledgment

This work was supported by the National Science Foundation through NSF award CCF-1320884.

References

[1] R. Layek and A. Datta, "Fault detection and intervention in biological feedback networks," *Journal of Biological Systems*, vol. 20, no. 4, pp. 441–453, 2012.

[2] T. Leifeld, Z. Zhang, and P. Zhang, "Fault detection for probabilistic boolean networks," in *Proceedings of the 2016 European Control Conference (ECC), Aalborg, Denmark*. IEEE, 2016, pp. 740–745.

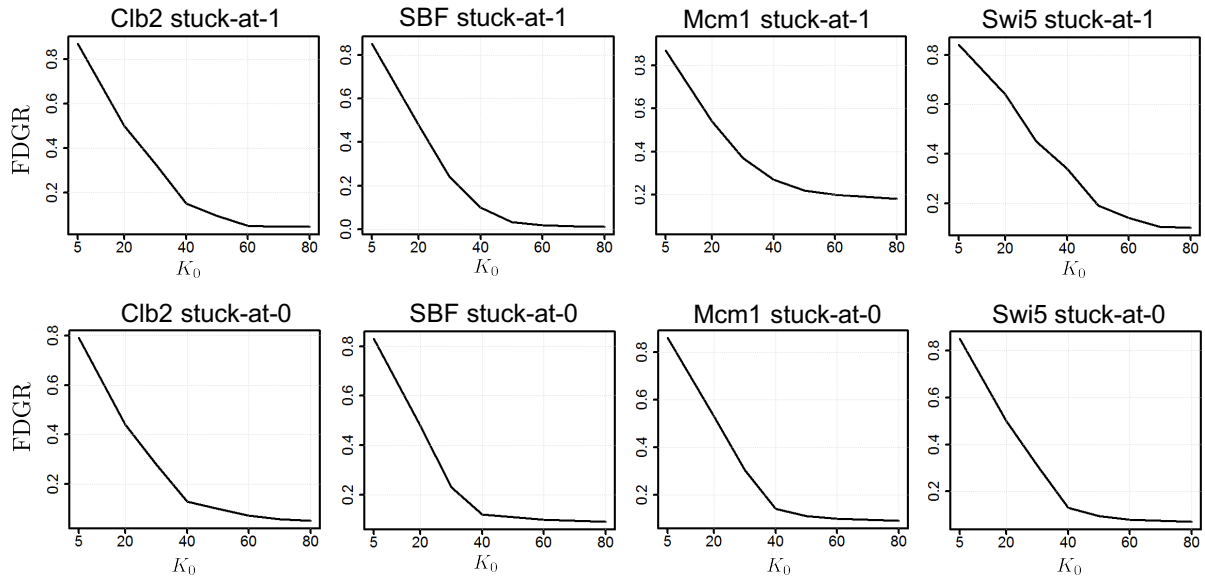

Fig. 3. False diagnosis rate (FDGR) for the LR method against the length K_0 of the initial fault-free sequence for systems identification.

TABLE I
PARAMETER VALUES USED IN THE NUMERICAL EXPERIMENTS.

Parameter	**Value**
Number of time series T	100
Length of time series K , IF method	400
Length of time series K , LR method	100
Fault time point k_0, IF method	$K_0 + 200$
Fault time point k_0, LR method	$K_0 + 50$
Number of particles N	20,000
Process noise intensity p	0.05
Sequencing depth s (reads)	1.0175 (1K-50K) 2.875 (50K-100K) 10.175 (250K-300K) 45.05 (1M-1M+50K)
Baseline expression μ_b	0.1
Differential expression mean $\delta_i, i = 1, \cdots, 14$	3
Inverse dispersion $\phi_i, i = 1, \cdots, 14$	5
Auto-correlation function lags L	15
Detection window length L_D	100
Certification threshold t_C	6
Certification window length L_C	10
LR thresholding parameter h	40
LR window parameter w_{LR}	10
APF-CPMLA-BKS stopping threshold ϵ	10^{-4}

[3] E. Fornasini and M. Valcher, "Fault detection analysis of boolean control networks," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2734–2739, 2015.

[4] S. Sridharan, R. Layek, A. Datta, and J. Venkatraj, "Boolean modeling and fault diagnosis in oxidative stress response," *BMC Genomics*, vol. 13, no. Suppl. 6, p. S4, 20.

[5] R. Weinberg, *The Biology of Cancer*. Princeton: Garland Science, 2006.

[6] S. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.

[7] I. Schmulevich, E. Dougherty, and W. Zhang, "From Boolean to probabilistic Boolean networks as models of genetic regulatory networks," *Proceedings of the IEEE*, vol. 90, pp. 1778–1792, 2002.

[8] G. Karlebach and R. Shamir, "Modelling and analysis of gene regulatory networks," *Nature Reviews Mollecular Cell Biology*, vol. 9, no. 10, pp. 770–780, 2008.

[9] I. Shmulevich and E. Dougherty, *Probabilistic Boolean Networks*. Philadelphia, PA: SIAM, 2009.

[10] U. Braga-Neto, "Optimal state estimation for Boolean dynamical systems," IEEE, pp. 1050–1054, 2011, proceedings of 45th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA.

[11] N. Berntenis and M. Ebeling, "Detection of attractors of large boolean networks via exhaustive enumeration of appropriate subspaces of the state space," *BMC Bioinformatics*, vol. 14, p. 361, 2013.

[12] P. Zhu and J. Han, "Asynchronous stochastic boolean networks as gene network models," *Journal of Computational Biology*, vol. 21, no. 10, pp. 771–783, 2014.

[13] M. Imani and U. M. Braga-Neto, "Maximum-likelihood adaptive filter for partially-observed Boolean dynamical systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 359–371, 2017.

[14] S. Bornholdt, "Boolean network models of cellular regulation: prospects and limitations," *J. R. Soc. Interface*, vol. 5, Suppl 1, pp. S85–S94, 2008.

[15] R. Albert and H. Othmer, "The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster," *Journal of theoretical biology*, vol. 223, no. 1, pp. 1–18, 2003.

[16] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang, "The yeast cell-cycle network is robustly designed," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 14, pp. 4781–6, 2004.

[17] A. Faure, A. Naldi, C. Chaouiya, and D. Thieffry, "Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle," *"Bionformatics"*, vol. 22, no. 14, pp. e124–e131, 2006.

[18] E. Batchelor, A. Loewer, and G. Lahav, "The ups and downs of p53: understanding protein dynamics in s ingle cells," *Nature Reviews Cancer*, vol. 9, pp. 371–377, 2009.

[19] M. Imani and U. Braga-Neto, "Optimal state estimation for Boolean dynamical systems using a Boolean Kalman smoother," in *Proceedings of the 3rd IEEE Global Conference on Signal and Information Processing (GlobalSIP'2015), Orlando, FL*. IEEE, 2015, pp. 972–976.

[20] A. Bahadorinejad and U. Braga-Neto, "Optimal fault detection and diagnosis in transcriptional circuits using next-generation sequencing," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 2, pp. 516–525, 2018.

TABLE II
PERFORMANCE EVALUATION RESULTS FOR FAR1=0, HOG=1

		LR Method			IF Method	
Fault class	Reads	FDR	FDGR	ATCD	FDR	ATCD
Clb2 stuck-at-1	1K-50K	0.12	0.17	10.21	0.15	11.51
	50K-100K	0.10	0.12	8.91	0.15	10.23
	250K-300K	0.06	0.08	6.69	0.12	8.12
	1M-1M+50K	0.04	0.06	5.01	0.09	7.38
SBF stuck-at-1	1K-50K	0.18	0.22	8.84	0.27	18.47
	50K-100K	0.17	0.19	7.63	0.23	16.34
	250K-300K	0.06	0.08	5.62	0.16	10.25
	1M-1M+50K	0.04	0.07	4.19	0.11	10.56
Mcm1 stuck-at-1	1K-50K	0.25	0.32	9.53	0.31	19.12
	50K-100K	0.19	0.24	8.33	0.24	18.02
	250K-300K	0.16	0.19	8.27	0.20	16.11
	1M-1M+50K	0.11	0.12	6.83	0.19	15.59
Swi5 stuck-at-1	1K-50K	0.31	0.32	13.78	0.32	17.93
	50K-100K	0.28	0.29	11.41	0.31	17.53
	250K-300K	0.09	0.17	9.33	0.24	16.71
	1M-1M+50K	0.08	0.10	8.01	0.14	13.74
Clb2 stuck-at-0	1K-50K	0.23	0.26	10.57	0.28	18.39
	50K-100K	0.20	0.25	9.78	0.32	14.92
	250K-300K	0.11	0.17	8.22	0.14	12.56
	1M-1M+50K	0.08	0.1	7.45	0.11	12.63
SBF stuck-at-0	1K-50K	0.27	0.28	15.42	0.38	18.96
	50K-100K	0.22	0.24	13.90	0.33	17.38
	250K-300K	0.16	0.17	9.31	0.29	14.49
	1M-1M+50K	0.12	0.11	9.04	0.17	13.62
Mcm1 stuck-at-0	1K-50K	0.12	0.19	7.32	0.16	14.39
	50K-100K	0.11	0.17	6.12	0.13	10.82
	250K-300K	0.09	0.11	5.9	0.11	9.36
	1M-1M+50K	0.06	0.10	4.92	0.09	9.29
Swi5 stuck-at-0	1K-50K	0.25	0.26	12.10	0.39	17.26
	50K-100K	0.21	0.25	8.12	0.28	13.52
	250K-300K	0.11	0.2	5.81	0.15	9.18
	1M-1M+50K	0.1	0.16	5.2	0.13	11.75

TABLE III
ESTIMATED PARAMETERS VALUES USING THE APF-CPMLA-BKS ALGORITHM WITH $K_0 = 50$.

	Estimated parameters															
reads	p	μ_b	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}	δ_{12}	δ_{13}	δ_{14}
1K-50K	0.064	0.12	3.25	2.84	2.60	3.14	3.23	2.71	2.80	3.20	2.64	2.81	2.92	3.13	3.33	2.49
50K-100K	0.065	0.12	3.20	2.95	2.60	3.09	3.10	2.74	2.93	3.04	2.74	2.92	2.95	3.08	3.10	2.48
250K-300K	0.064	0.10	3.14	2.97	2.63	3.09	3.11	2.83	2.95	3.00	2.69	2.92	2.99	3.024	3.08	2.44
1M-1M+50K	0.056	0.10	3.05	2.98	2.85	3.05	3.10	2.89	3.03	3.00	2.88	2.95	3.00	3.01	3.00	2.89

[21] M. Imani and U. Braga-Neto, "Particle filters for partially-observed Boolean dynamical systems," *Automatica*, vol. 87, pp. 238–250, 2018.

[22] ——, "Optimal finite-horizon sensor selection for Boolean Kalman filter," in *2017 51th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2017.

[23] M. Imani and U. M. Braga-Neto, "Point-based methodology to monitor

and control gene regulatory networks via noisy measurements," *IEEE Transactions on Control Systems Technology*, 2018.
[24] ——, "Control of gene regulatory networks with noisy measurements and uncertain inputs," *IEEE Transactions on Control of Network Systems*, 2018, dOI: 10.1109/TCNS.2017.2746341.
[25] M. Imani and U. Braga-Neto, "Multiple model adaptive controller for partially-observed Boolean dynamical systems," in *Proceedings of the 2017 American Control Conference (ACC'2017), Seattle, WA*. IEEE, 2017, pp. 1103–1108.
[26] L. D. Mcclenny, M. Imani, and U. M. Braga-Neto, "BoolFilter: an R package for estimation and identification of partially-observed Boolean dynamical systems," *BMC bioinformatics*, vol. 18, no. 1, p. 519, 2017.
[27] D. Himmelblau, *Fault Detection and Diagnosis in Chemical and Petrochemical P rocesses*. New York: Elsevier, 1978.
[28] J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*. New York: Marcel Dekker, 1998.
[29] J. Chen and R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Boston: Kluwer, 1999.
[30] J. Prakash, S. Patwardhan, and S. Narasimhan, "A supervisory approach to fault-tolerant control of linear multivariable systems," *Industrial and Engineering Chemical Research*, vol. 41, pp. 2270–2281, 2002.
[31] R. Isermann, *Fault-Diagnosis Systems*. Heildelberg: Springer, 2006.
[32] K. Villeza, B. Srinivasanb, R. Rengaswamyb, S. Narasimhanc, and V. Venkatasubramanian, "Kalman-based strategies for fault detection and identification (fdi): Extensions and critical evaluation for a buffer tank system," *Computers and Chemical Engineering*, vol. 35, pp. 806–816, 2010.
[33] L. Chiang, E. Russell, and R. Braatz, *Fault detection and diagnosis in industrial systems*. New York: Springer, 2013.
[34] A. SHIRYAEV, "The problem of the most rapid detection of a disturbance in a stationary process," *Soviet Math. Dokl.*, 1961.
[35] G. Lorden, "Procedures for reacting to a change in distribution," in *The Annals of Mathematical Statistics*, 1971.
[36] E. PAGE, "Continuous inspection schemes," *Biometrika*, 1954.
[37] P. Maybeck and P. Hanlon, "Performance enhancement of a multiple model adaptive estimator," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 4, pp. 1240–1253, 1995.
[38] A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, and B. Wold, "Mapping and quantifying mammalian transcriptomes by RNA-seq," *Nature methods*, vol. 5, no. 7, pp. 621–628, 2008.
[39] T. J. Hardcastle and K. A. Kelly, "bayseq: empirical bayesian methods for identifying differential expression in sequence count data," *BMC bioinformatics*, vol. 11, no. 1, p. 422, 2010.
[40] S. Anders and W. Huber, "Differential expression analysis for sequence count data," *Genome biol*, vol. 11, no. 10, p. R106, 2010.
[41] M. Imani and U. Braga-Neto, "Point-based value iteration for partially-observed Boolean dynamical systems with finite observation space," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 4208–4213.
[42] N. Ghaffari, M. R. Yousefi, C. D. Johnson, I. Ivanov, and E. R. Dougherty, "Modeling the next generation sequencing sample processing pipeline for the purposes of classification," *BMC bioinformatics*, vol. 14, no. 1, p. 307, 2013.
[43] D. W. J. Li, I. Johnstone, and R. Tibshirani, "Normalization, testing, and false discovery rate estimation for RNA-sequencing data," *Cellular and molecular life science*, 2012.
[44] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.
[45] A. Gelb, *Applied Optimal Estimation*. MIT Press, 1974.
[46] K. Villeza, B. Srinivasanb, R. Rengaswamyb, S. Narasimhanc, and V. Venkatasubramanian, "Kalman-based strategies for fault detection and identification (fdi): Extensions and critical evaluation for a buffer tank system," *Computers and Chemical Engineering*, 2010.
[47] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Inc, 1993.
[48] E. Radmaneshfar and M. Thiel, "Recovery from stress–a cell cycle perspective," *Journal of computational interdisciplinary sciences*, vol. 3, no. 1-2, p. 33, 2012.

Arghavan Bahadorinejad received a Ph.D. degree in Electrical and Computer Engineering from Texas A&M University, College Station, TX in 2017. She is currently with Apple Inc., Cupertino, CA. Her Ph.D. research included the study of discrete-time Boolean dynamical systems and fault detection, with applications in genomic signal processing.

Mahdi Imani received his B.Sc. degree in Mechanical Engineering and his M.Sc. degree in Electrical Engineering, both from University of Tehran in 2012 and 2014. He is currently a Ph.D. student at the Department of Electrical and Computer Engineering of Texas A&M University, College Station, TX. His research interests include machine learning, control theory and signal processing.

Ulisses M. Braga-Neto lisses M. Braga-Netolisses M. Braga-NetoU is an Associate Professor at the Department of Electrical and Computer En- gineering and a member of the Center for Bioinformatics and Genomic Systems Engineering at Texas A&M University, College Station, TX. He holds a Ph.D. degree in Electrical and Computer Engineering from The Johns Hopkins University, Baltimore, MD and has held post-doctoral positions at the University of Texas M.D. Anderson Cancer Center, Houston, TX and at the Oswaldo Cruz Foundation, Recife, Brazil. His research interests include Pattern Recognition and Statistical Signal Processing. Dr. Braga-Neto is a Senior Member of the IEEE. He is author of the textbook Error Estimation for Pattern Recognition (IEEE-Wiley, 2015) and has received the NSF CAREER Award for his work in this area.