# Optimal State Estimation for Boolean Dynamical Systems using a Boolean Kalman Smoother

*The 3rd IEEE Global Conference on Signal & Information Processing*

**Mahdi Imani and Ulisses Braga-Neto**

Department of Electrical Engineering
Texas A&M University
College Station, Texas 77843

December 16, 2015

# Outline I

# Introduction

- Boolean networks have emerged as an effective model of the dynamical behavior of regulatory circuits consisting of bi-stable components.
- In the Boolean network model, the transcriptional state of each gene is represented by 0 (OFF) or 1 (ON), and the relationship among genes is described by logical gates updated and observed at discrete time intervals.
- This model has been successful in accurately modeling the dynamics of the cell cycle in the Drosophila fruit fly, in the Saccharomyces cerevisiae yeast, as well as the mammalian cell cycle.

# Outline I

# Partially-Observable Boolean Dynamical Systems

- *Boolean State Transition Model:* there is uncertainty in state transition. The sequence of state vectors $\{\mathbf{X}_k; k = 0, 1, \ldots\}$ is a Markov stochastic process, called the *state process*, specified by

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

  $\mathbf{u}_{k-1}$ and $\mathbf{f}$ are the input and network function, respectively, whereas $\{\mathbf{n}_k; k = 1, 2, \ldots\}$ is a white noise process.

- *Observation 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:

$$\mathbf{Y}_k \; = \; \mathbf{h}\,(\mathbf{X}_k \,, \, \mathbf{v}_k)\,, \tag{2}$$

  where $\mathbf{v}_k$ is the observation noise.

# Outline I

# Boolean Kalman Filter

The Boolean Kalman Filter (BKF) is the recursive minimum mean-square error (MMSE) state estimator $\hat{\mathbf{X}}_k = h(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$ of the state $\mathbf{X}_k$, according to the (conditional) mean-square error (MSE):

$$\mathrm{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) = E\left[ ||\hat{\mathbf{X}}_k - \mathbf{X}_k||^2 \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1 \right] \quad (3)$$

## Theorem

**(Boolean Kalman Filter.)** *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* $k$, *is given by*

$$\hat{\mathbf{X}}_k = \overline{E\left[\mathbf{X}_k \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right]}, \quad (4)$$

*where* $\overline{\mathbf{v}}(i) = I_{\mathbf{v}(i) > 1/2}$ *for* $i = 1, \ldots, d$.

- Define the following distribution vectors of length $2^d$:

$$\begin{aligned} \mathbf{\Pi}_{k|k}(i) &= P\left(\mathbf{X}_k = \mathbf{x}^i \mid \mathbf{Y}_k, \ldots, \mathbf{Y}_1\right) , \\ \mathbf{\Pi}_{k|k-1}(i) &= P\left(\mathbf{X}_k = \mathbf{x}^i \mid \mathbf{Y}_{k-1}, \ldots, \mathbf{Y}_1\right) , \\ \mathbf{\Delta}_{k|k}(i) &= P\left(\mathbf{Y}_{k+1}, \ldots, \mathbf{Y}_T \mid \mathbf{X}_k = \mathbf{x}^i\right) , \\ \mathbf{\Delta}_{k|k-1}(i) &= P\left(\mathbf{Y}_k, \ldots, \mathbf{Y}_T \mid \mathbf{X}_k = \mathbf{x}^i\right) , \end{aligned} \tag{5}$$

- Prediction Matrix:

$$\begin{aligned} (M_k)_{ij} &= P(\mathbf{X}_k = \mathbf{x}^i \mid \mathbf{X}_{k-1} = \mathbf{x}^j) \\ &= P\left(\mathbf{n}_k = \mathbf{x}^i \oplus \mathbf{f}(\mathbf{x}^j, \mathbf{u}_{k-1})\right) , \end{aligned} \tag{6}$$

- Update Matrix:

$$(T_k)_{jj} = P\left(\mathbf{Y}_k \mid \mathbf{X}_k = \mathbf{x}^j\right) , \tag{7}$$

- Boolean States: $A = [\mathbf{x}^1 \ldots \mathbf{x}^{2^d}]$

## Boolean Kalman Filter

1) *Initialization Step: The initial PDV is given by* $\boldsymbol{\Pi}_{0|0}(i) = P\left(\mathbf{X}_0 = \mathbf{x}^i\right)$, *for* $i = 1, \ldots, 2^d$.

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

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

$$\boldsymbol{\Pi}_{k|k-1} = M_k \, \boldsymbol{\Pi}_{k-1|k-1} \, .$$

3) *Update Step: Given the current observation* $\mathbf{Y}_k = \mathbf{y}_k$, *let* $\boldsymbol{\beta}_k = T_k(\mathbf{y}_k) \, \boldsymbol{\Pi}_{k|k-1}$ .

*The* updated *PDV* $\boldsymbol{\Pi}_{k|k}$ *is obtained by normalizing* $\boldsymbol{\beta}_k$: $\boldsymbol{\Pi}_{k|k} = \dfrac{\boldsymbol{\beta}_k}{||\boldsymbol{\beta}_k||_1}$

4) *MMSE Estimator Computation Step: The MMSE estimator is given by* $\hat{\mathbf{X}}_k = \overline{A\boldsymbol{\Pi}_{k|k}}$

*with optimal conditional MSE:* $\text{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) = ||\min\{A\boldsymbol{\Pi}_{k|k}, (A\boldsymbol{\Pi}_{k|k})^c\}||_1$ .

# Outline I

# Boolean Kalman Smoother

The Boolean Kalman Smoother (BKS) is the minimum mean-square error (MMSE) state estimator $\hat{\mathbf{X}}_S = g(\mathbf{Y}_1, \ldots, \mathbf{Y}_T)$ of the state $\mathbf{X}_S$, where $1 < S < T$, according to the (conditional) mean-square error (MSE):

$$\mathrm{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_T) = E\left[ ||\hat{\mathbf{X}}_S - \mathbf{X}_S||^2 \mid \mathbf{Y}_T, \ldots, \mathbf{Y}_1 \right] \tag{8}$$

## Theorem

**(Boolean Kalman Smoother.)** *The optimal minimum MSE estimator $\hat{\mathbf{X}}_S$ of the state $\mathbf{X}_S$ given the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_T$, where $1 < S < T$, is given by*

$$\hat{\mathbf{X}}_S = \overline{E\left[\mathbf{X}_S \mid \mathbf{Y}_T, \ldots, \mathbf{Y}_1\right]}, \tag{9}$$

*where $\overline{\mathbf{v}}(i) = I_{\mathbf{v}(i) > 1/2}$ for $i = 1, \ldots, d$.*

# Boolean Kalman Smoother

**Boolean Kalman Smoother**

*Forward Estimator:*

1) *Initialization Step:* $\mathbf{\Pi}_{1|0} = M_1 \mathbf{\Pi}_{0|0}$.

*For* $k = 1, 2, \ldots, S-1$, *do:*

2) *Update Step:* $\boldsymbol{\beta}_k = T_k(\mathbf{y}_k)\, \mathbf{\Pi}_{k|k-1}$.
3) *Normalization Step:* $\mathbf{\Pi}_{k|k} = \boldsymbol{\beta}_k / ||\boldsymbol{\beta}_k||_1$.
4) *Prediction Step:* $\mathbf{\Pi}_{k+1|k} = M_{k+1}\, \mathbf{\Pi}_{k|k}$.

*Backward Estimator:*

1) *Initialization Step:* $\mathbf{\Delta}_{T|T-1} = T_T(\mathbf{y}_T) \mathbf{1}_{d\times 1}$.

*For* $k = T-1, T-2, \ldots, S$, *do:*

2) *Prediction Step:* $\mathbf{\Delta}_{k|k} = {M_{k+1}}^T\, \mathbf{\Delta}_{k+1|k}$.
3) *Update Step:* $\mathbf{\Delta}_{k|k-1} = T_k(\mathbf{y}_k)\, \mathbf{\Delta}_{k|k}$.

*Smoothed Distribution Vector:*

$$\mathbf{\Pi}_{S|T} = \frac{\mathbf{\Pi}_{S|S-1} \bullet \mathbf{\Delta}_{S|S-1}}{||\mathbf{\Pi}_{S|S-1} \bullet \mathbf{\Delta}_{S|S-1}||_1},$$

*where "$\bullet$" denotes component-wise vector multiplication.*

*MMSE Estimator:*

*The MMSE estimator is given by:*

$$\hat{\mathbf{X}}_S = \overline{A\mathbf{\Pi}_{S|T}}$$

*with optimal conditional MSE*

$$\mathrm{MSE}(\mathbf{Y}_1, \ldots, \mathbf{Y}_T) = ||\min\{A\mathbf{\Pi}_{S|T}, (A\mathbf{\Pi}_{S|T})^c\}||_1 .$$

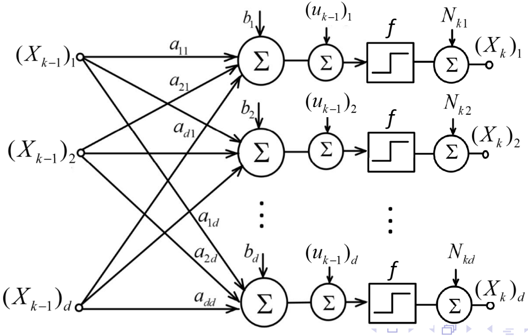
# Outline I

- Parametric State Transition Model:

$$f_i(\mathbf{X}_{k-1}, \mathbf{u}_{k-1}) = \begin{cases} 1, & \sum_j a_{ij}(\mathbf{X}_{k-1})_j + b_i + (\mathbf{u}_{k-1})_i > 0 \\ 0, & \sum_j a_{ij}(\mathbf{X}_{k-1})_j + b_i + (\mathbf{u}_{k-1})_i < 0 \end{cases} \quad (10)$$

$a_{ij} = +1$ (Positive Regulation), $a_{ij} = -1$ (Negative Regulation), $a_{ij} = 0$ (No Regulation)

$b_i = +1/2$ (Positively Biased), $b_i = -1/2$ (Negatively Biased)



- *RNA-Seq Observation Model:*
  1. RNA-seq Data

     In this study, we choose to use a Poisson model for the number of reads for each transcript:

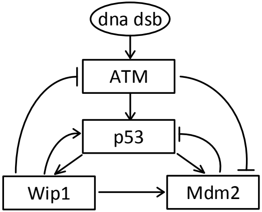
$$P(Y_{ki} = m \mid \lambda_{ki}) = e^{-\lambda_{ki}} \frac{\lambda_{ki}^m}{m!}, \quad m = 0, 1, \ldots \tag{11}$$

$\lambda_{ki}$ is the mean read count of transcript $i$ at time $k$

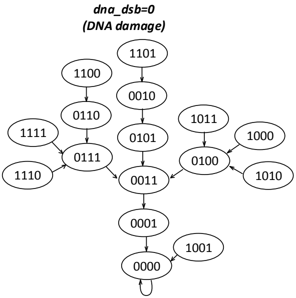
$$\begin{aligned} \log(\lambda_{ki}) &= \log(s) + \mu_b\,, && \text{if } X_{ki} = 0\,, \\ \log(\lambda_{ki}) &= \log(s) + \mu_b + \delta_i\,, && \text{if } X_{ki} = 1\,. \end{aligned} \tag{12}$$

$s$: sequencing depth,
$\mu_b > 0$: Baseline expression,
$\delta_i > 0$: Differential expression .

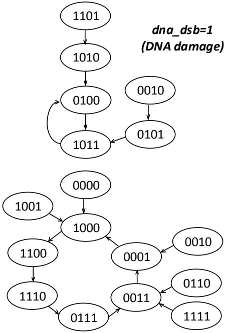
# Case Study: p53-MDM2 Negative Feedback Loop Pathway




(a)



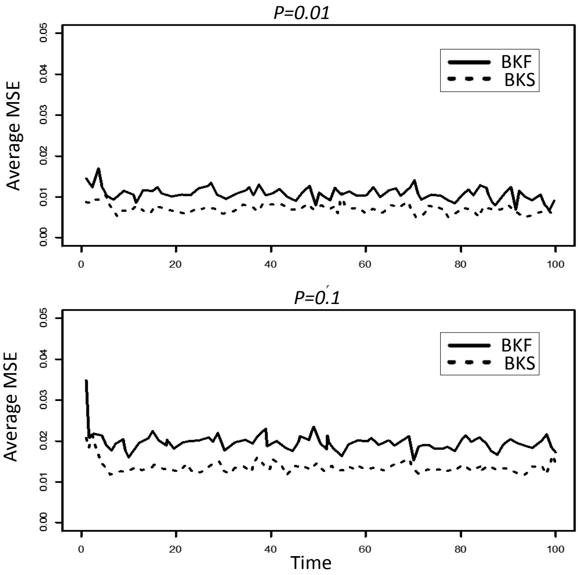

(b)




(c)

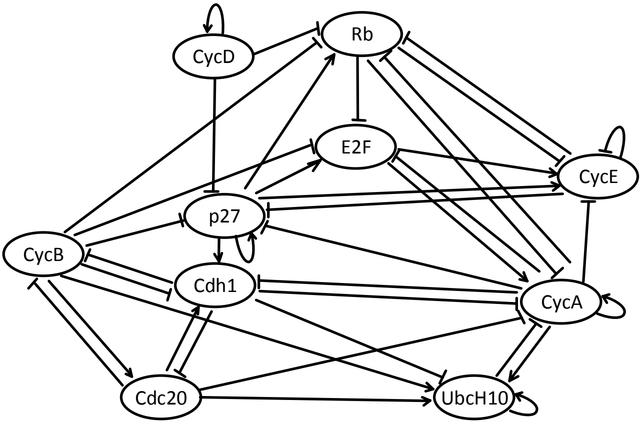
| | | *dna_dsb = 0* | | *dna_dsb = 1* | |
|---|---|---|---|---|---|
| **Noise** | **Reads** | **BKF** | **BKS** | **BKF** | **BKS** |
| $p = 0.01$ | 1K-50K | 0.95 | 0.97 | 0.89 | 0.92 |
| | 500K-550K | 0.98 | 0.99 | 0.92 | 0.95 |
| $p = 0.1$ | 1K-50K | 0.86 | 0.91 | 0.82 | 0.86 |
| | 500K-550K | 0.88 | 0.94 | 0.84 | 0.88 |

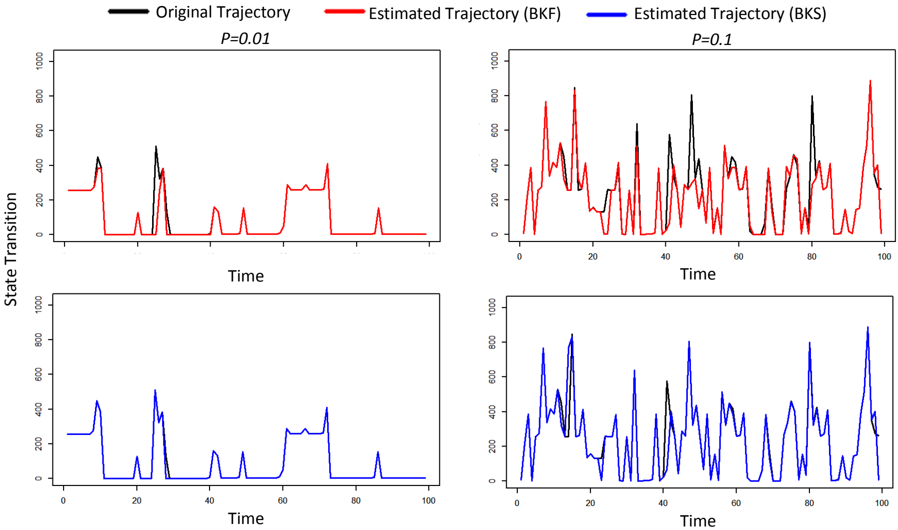
Average Performance of BKF and BKS.




Average MSE of BKF and BKS over 1000 runs.

# Case Study: Cell Cycle Regulatory Network






Estimated Trajectories by BKF and BKS.

# Outline I

# Conclusions and Future Work

- We proposed a method for state estimation for Boolean dynamical system observed though a single time series of noisy measurements given the entire history of observations.
- Future work includes:
  - Dealing with the network inference problem in the presence of batch data.
  - Developing methods for discrete, continuous and mixed parameter estimation.
  - Deriving efficient methods for large networks exploring sparsity, for both state and parameter estimation.

# Thank you

# Q&A