Sequential Change Detection:
an Overview

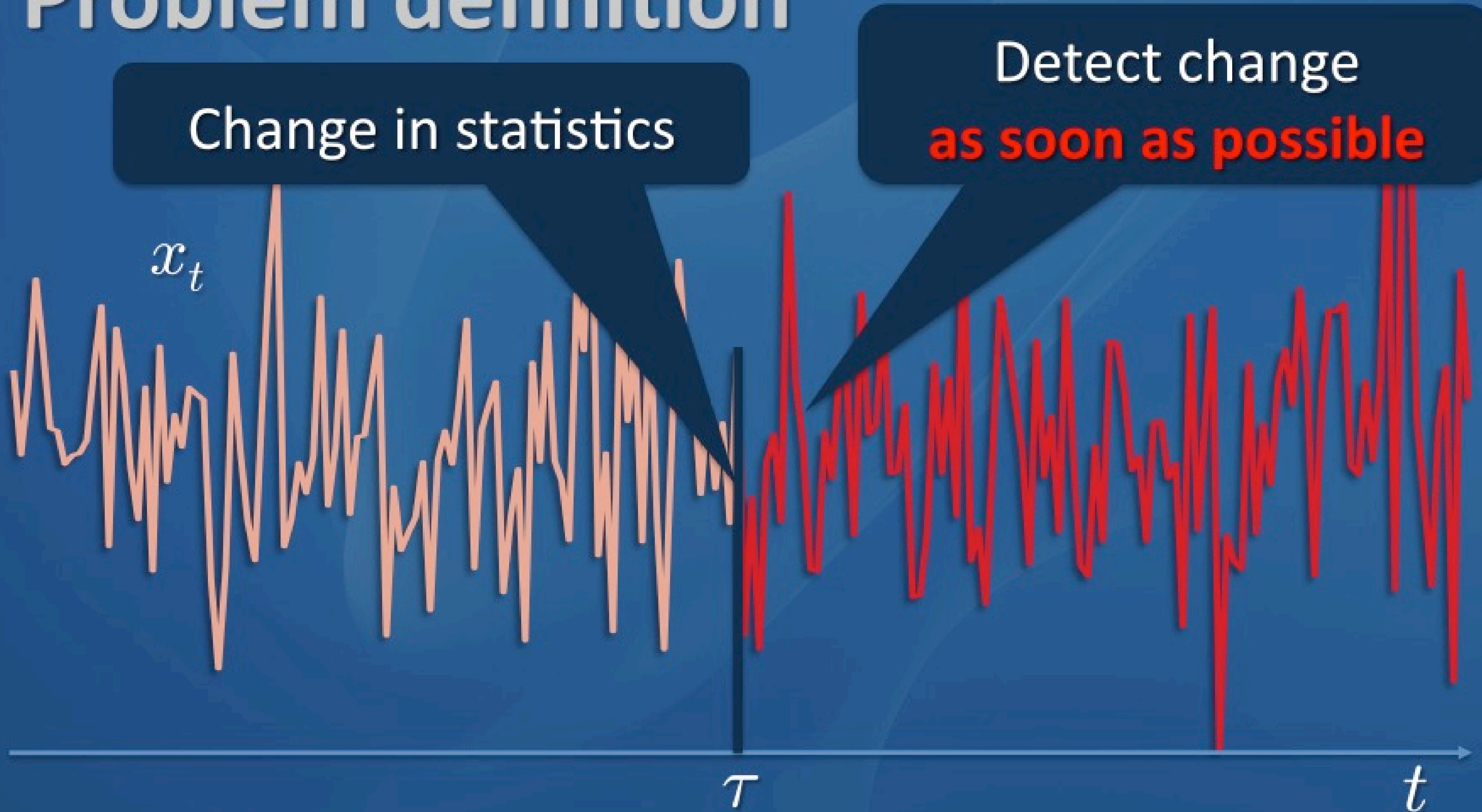
(Quickest Detection, Sequential Anomaly Detection)

George V. Moustakides
University of Patras, Greece

Outline

- Problem definition: Change generation mechanisms and detectors
- Formulations involving expected delays
- Formulations involving hard delay constraints
- Decentralized detection (sensor networks)
- Intrusion detection in wireless networks

Problem definition



Specify: a) Detector form

b) Change generation mechanisms

Applications

Quality monitoring of manufacturing process (1930's)
Biomedical Engineering
Electronic Communications
Econometrics
Seismology
Speech & Image Processing (segmentation)
Vibration monitoring (Structural health monitoring)
Security monitoring (fraud detection)
Spectrum monitoring
Scene monitoring
Network monitoring (router failures, **anomaly** detection)
Epidemic detection .....

CUSUM: 2,280 hits in 2013. Google Scholar

We observe **sequentially** a process $\{x_t\}$ that has the following statistical properties

$$x_t \sim \begin{cases} f_0 & \text{for } 0 < t \leq \tau \\ f_1 & \text{for } t > \tau \end{cases}$$

Changetime is unknown!!!

Detect occurrence of τ as soon as possible

At every time t consult available data: x_1 ,…, x_t , x_{t+1}

- ~~Change did not take place before t~~
  Continue sampling
- ~~Change took place before t~~
  Stop sampling!

Stopping time

Sequential Detector ←→ Stopping time

Stopping times

We observe sequentially a process $\{x_t\}$

A random time $T \in \{0, 1, 2, \ldots\}$ is called a stopping time adapted to $\{x_t\}$ when the event $\{T = t\}$ depends only on $\{x_1, \ldots, x_t\}$

Optimal Stopping Theory

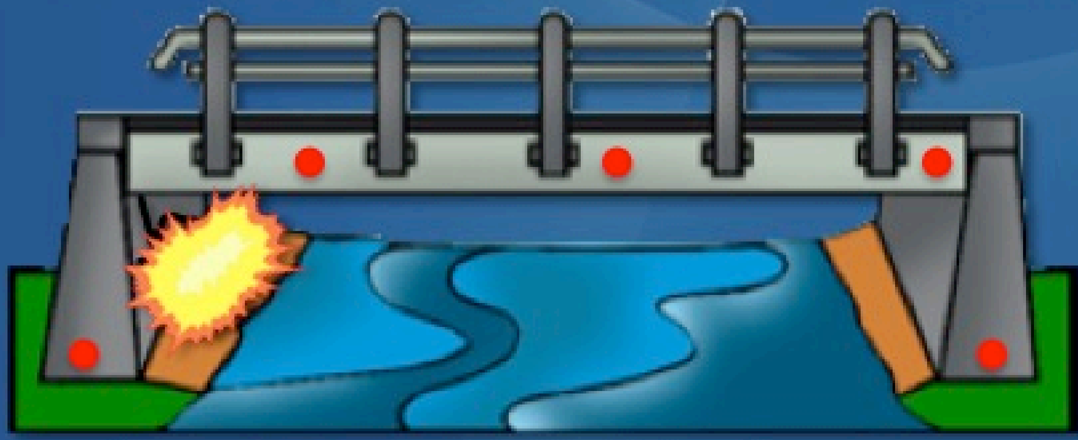
For $\{\phi_t(x)\}, \{\alpha_t(x)\}$ deterministic functions

Optimize

$$\mathsf{E}[\phi_T(x_T)] \quad \text{or } \mathsf{E}\left[\sum_{t=0}^{T-1} \alpha_t(x_t) + \phi_T(x_T)\right]$$

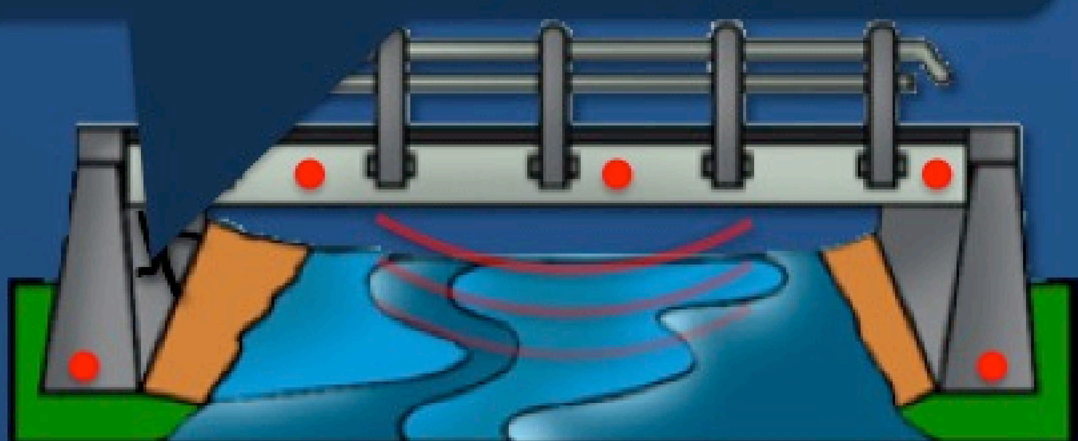
Change generation mechanisms

Structural health monitoring



Change mechanism **independent** from data

Amplitude of oscillations overly large



Change mechanism **dependent** on data

Formulations with expected delays

Pre-change : P_0 (E_0); Post-change : P_1 (E_1);

We are looking for a stopping time T.

General criterion:

$$J(T) = \mathsf{E}_1[T - \tau \mid T > \tau]$$

Change mechanism independent from data

Shiryaev (1963): τ is random with **known** prior.

Exponential prior: $\mathsf{P}(\tau = t) = p(1-p)^t$

$$\inf_T J(T) \quad \text{subject to} : \mathsf{P}_0(T \leq \tau) \leq \alpha$$

Discrete time: i.i.d. data before and after the change with pdfs f_0,f_1.

Define the statistic : $S_t = (S_{t-1} + 1)\dfrac{f_1(x_t)}{f_0(x_t)(1-p)}$

$$T_{\mathsf{S}} = \inf\{t > 0 : S_t \geq \nu\}$$

Threshold $\nu > 0$ such that the false alarm constraint is satisfied with equality.

In continuous time when $\{x_t\}$ is a Brownian motion with constant drift before and after the change.
In continuous time when $\{x_t\}$ is Poisson with constant rate before and after the change.

Time variation? Dependence? Multiple pre- and/or post-change possibilities?

$$J(T) = \mathsf{E}_1[T - \tau \mid T > \tau]$$

Changetime τ is random with **unknown** prior.

Pollak (1985): Follow a **worst-case analysis** for the prior.

$$J_{\mathsf{P}}(T) = \sup_{\text{all priors}} \mathsf{E}_1[T - \tau \mid T > \tau]$$

We can show:

$$J_{\mathsf{P}}(T) = \sup_{t \geq 0} \mathsf{E}_1[T - t \mid T > t]$$

$$\inf_T J_{\mathsf{P}}(T) \quad \text{subject to : } \mathsf{E}_0[T] \geq \gamma$$

Discrete time: i.i.d. data before and after the change with pdfs f_0,f_1.

Compute recursively the following statistic:

$$S_t = (S_{t-1} + 1)\frac{f_1(x_t)}{f_0(x_t)}$$

$$T_{\mathsf{P}} = \inf\{t > 0 : S_t \geq \nu\}$$

Pollak (1985): S_0 if specially designed, then

$$[J_{\mathsf{P}}(T_{\mathsf{P}}) - \inf_T J_{\mathsf{P}}(T)] \to 0; \quad \text{as } \gamma \to \infty$$

Exact optimality? Tartakovsky (2012) counterexample.

Continuous-time? Time variation? Dependence? Multiple pre- and/or post-change possibilities?

Change mechanism dependent on data.

Lorden (1971): τ **unknown** dependence.

Follow a worst-case analysis.

$$J_{\mathsf{L}}(T) = \sup_{\text{data dependent } \tau} \mathsf{E}_1[T - \tau \mid T > \tau]$$

$$J_{\mathsf{L}}(T) = \sup_{t \geq 0} \sup_{x_1,\ldots,x_t} \mathsf{E}_1[(T-t)^+ \mid x_1, \ldots, x_t]$$

$$\inf_{T} J_{\mathsf{L}}(T) \quad \text{subject to} : \mathsf{E}_0[T] \geq \gamma$$

CUSUM stopping time:

$$u_t = \log\left(\frac{f_1(x_1,\ldots,x_t)}{f_0(x_1,\ldots,x_t)}\right); \quad \text{running LLR}$$

$$m_t = \inf_{0<s\leq t} u_s; \quad \text{running minimum}$$

$$S_t = u_t - m_{t-1}; \quad \text{CUSUM statistic}$$

$$T_{\mathsf{C}} = \inf\{t>0 : S_t \geq \nu\}; \quad \text{CUSUM stop. time}$$

For i.i.d. $S_t = (S_{t-1})^+ + \log\left(\frac{f_1(x_t)}{f_0(x_t)}\right)$


ν
ν
ν
u_t
m_t
T_C
ML estimate of τ

Discrete time: i.i.d. before and after the change

Lorden (1971) asymptotic optimality (order-1).

Moustakides (1986) strict optimality.

Continuous time

Shiryaev (1996), Beibel (1996) strict optimality for Brownian Motion with constant drifts before and after.

Moustakides (2004) strict optimality for Ito processes

Moustakides (under review) strict optimality for Poisson processes.

Time variation? Dependence? Multiple pre- and/or post-change possibilities?

Formulations with hard constraints

$$J(T) = \mathsf{E}_1[\underbrace{T-\tau}\mid T>\tau]$$

Detection delay can be **arbitrarily large!**

Several applications require detection delay **at most m**.

$$\tau < T \leq \tau + m$$

If $\tau + m < T,$ this is regarded as failure.

$$\mathcal{J}(T) = \mathsf{P}_1(T \le \tau + m \mid T > \tau)$$

Interested in detection probability

Change mechanism independent from data

τ random with known prior. (Shiryaev-like)

$$\sup_T \mathcal{J}(T) \text{ subject to : } \mathsf{P}_0(T \le \tau) \le \alpha$$

τ random with unknown prior. (Pollak-like)

$$\mathcal{J}_\mathsf{P}(T) = \inf_{t \ge 0} \mathsf{P}_1(T \le t + m \mid T > t)$$

$$\sup_T \mathcal{J}_\mathsf{P}(T) \quad \text{subject to : } \mathsf{E}_0[T] \ge \gamma$$

Change mechanism dependent on data.

τ unknown dependence (Lorden-like)

$$\mathcal{J}_{\mathsf{L}}(T) = \inf_{t\geq 0} \inf_{x_1,\ldots,x_t} \mathsf{P}_1(T \leq t+m \mid x_1,\ldots,x_t)$$

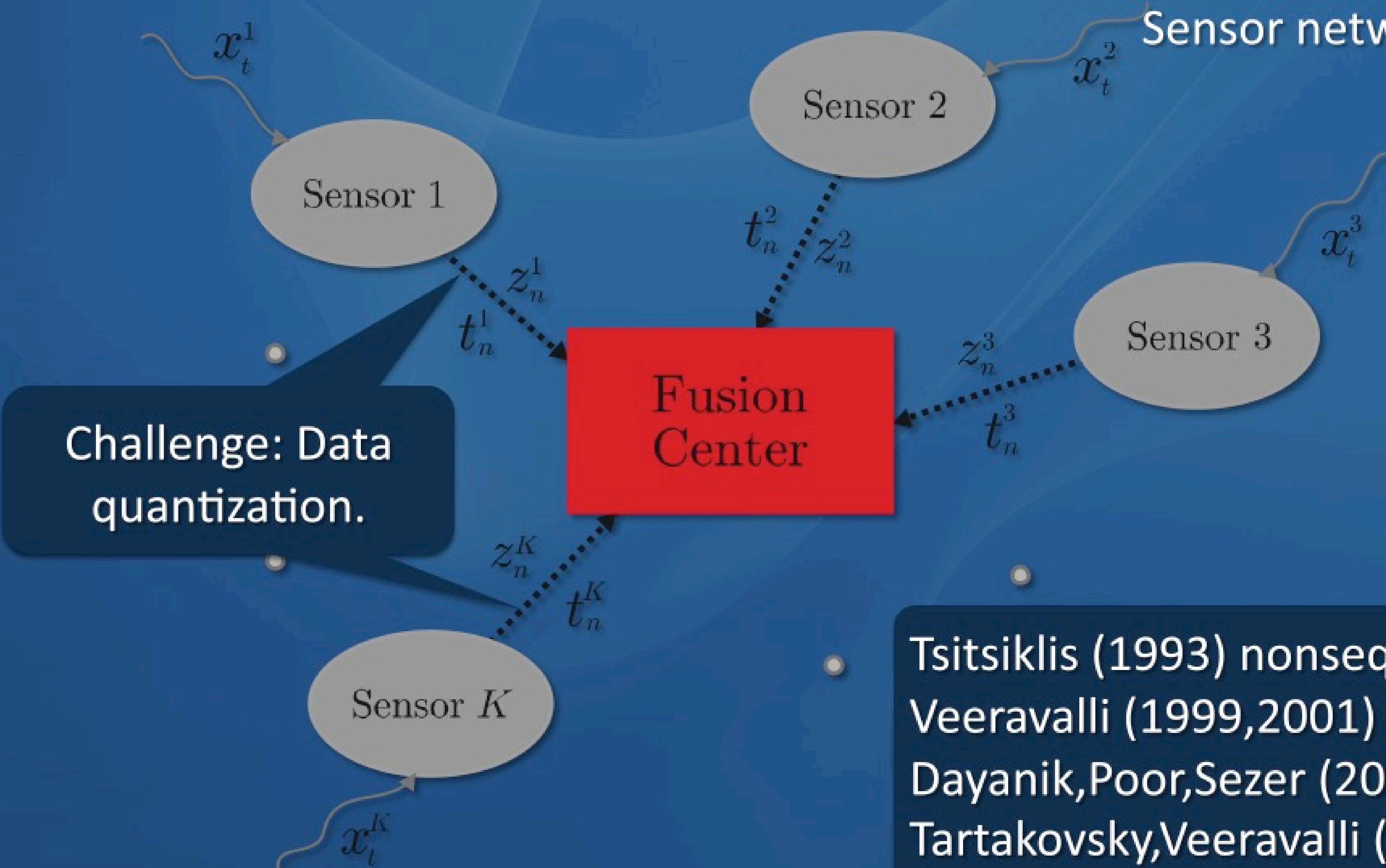
$$\sup_{T} \mathcal{J}_{\mathsf{L}}(T) \quad \text{subject to : } \mathsf{E}_0[T] \geq \gamma$$

Exact solution only for $m=1$ ($T=\tau+1$, i.e. detect the change with the first sample under the alternative regime).

$$T_{\mathsf{Sh}} = \inf\left\{t > 0 : \frac{f_1(x_t)}{f_0(x_t)} \geq \nu\right\}$$

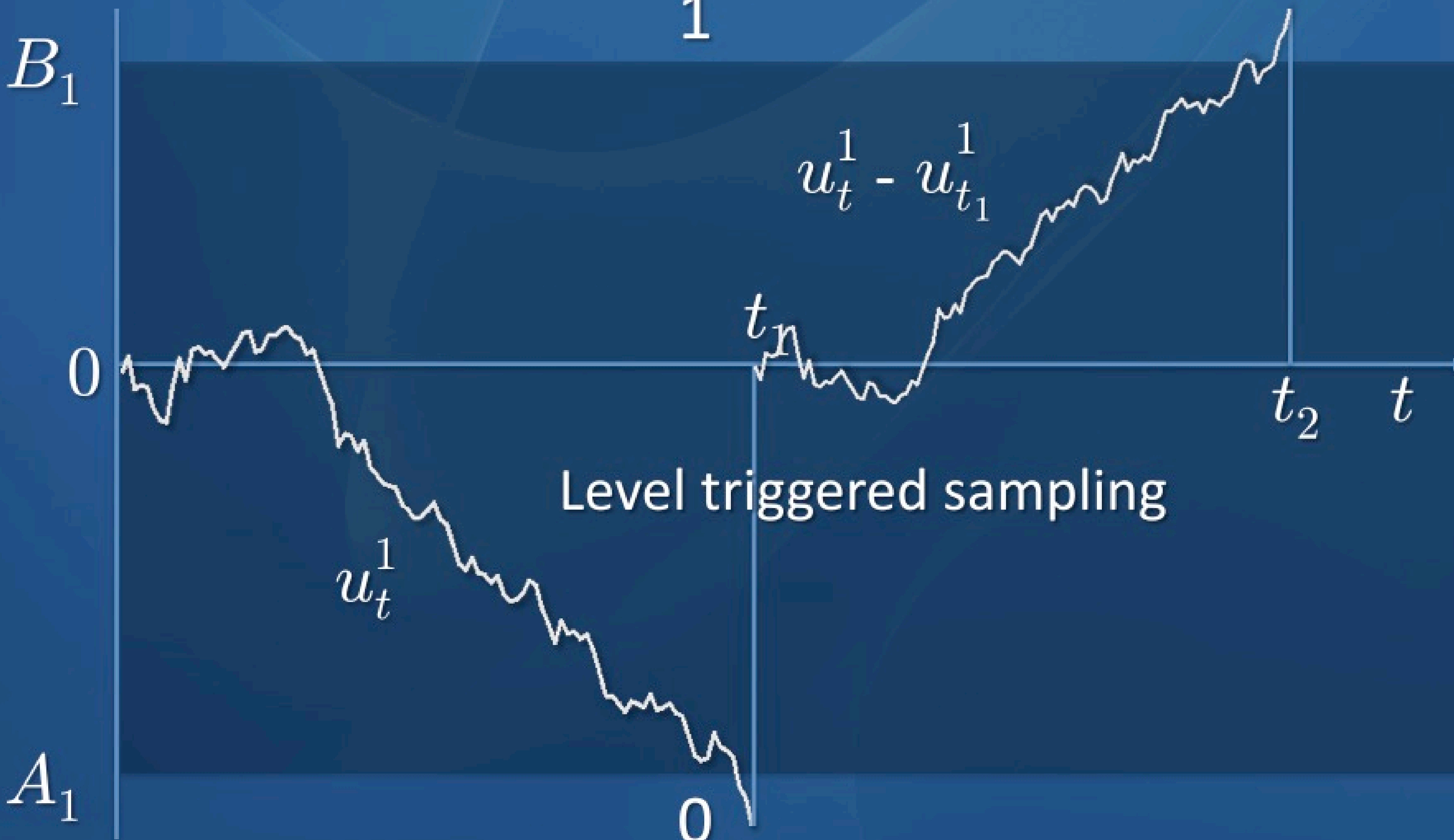
Shewhart (1931). Optimality: Bojdecki (1979); Pollak and Krieger (2013); Moustakides (2014).

Decentralized detection



Tsitsiklis (1993) nonsequen.
Veeravalli (1999,2001)
Dayanik,Poor,Sezer (2008)
Tartakovsky,Veeravalli (2008)

Fellouris-Moustakides (under review)



If more than 1 bits, **quantize overshoot!**

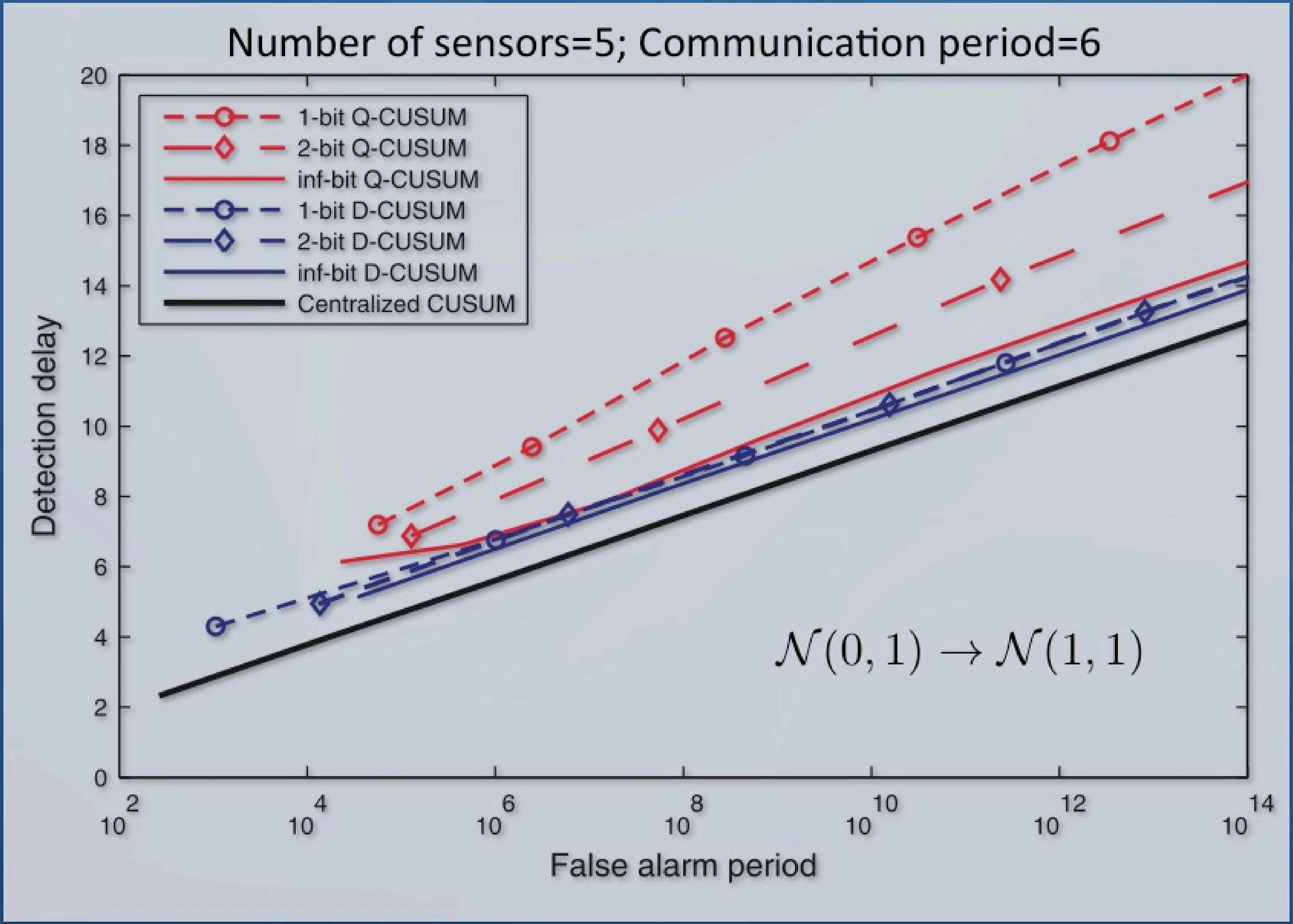
Communication with Fusion Center is:

- at random times
- asynchronous
- control over **average** communication rate with A_i,B_i

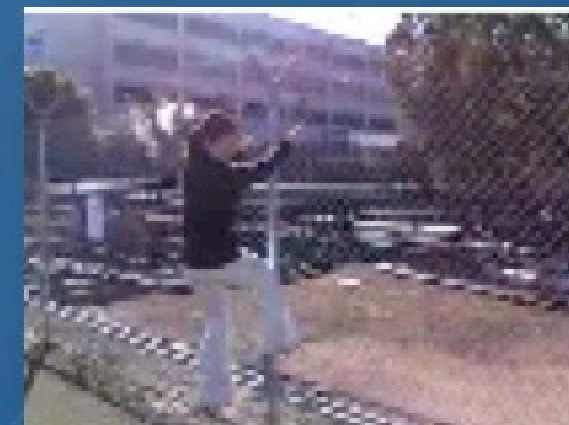
If sensor i sends a bit at time t, the Fusion Center updates an estimate of the global log-likelihood ratio:

$$\hat{u}_t = \begin{cases} \hat{u}_{t-} + B_i & \text{if bit is } 1 \\ \hat{u}_{t-} + A_i & \text{if bit is } 0 \end{cases}$$

and performs a CUSUM test using the estimate of the global log-likelihood ratio.


Number of sensors=5; Communication period=6
1-bit Q-CUSUM
2-bit Q-CUSUM
inf-bit Q-CUSUM
1-bit D-CUSUM
2-bit D-CUSUM
inf-bit D-CUSUM
Centralized CUSUM
Detection delay
False alarm period
$\mathcal{N}(0,1) \to \mathcal{N}(1,1)$
0
2
4
6
8
10
12
14
16
18
20
10^2
10^4
10^6
10^8
10^{10}
10^{12}
10^{14}

Smart Fence: Chraim and Pister, U of California, Berkeley (2013)



Long-term deployment setup at the Chevron-Richmond refinery. The result of this test was a detection rate of 100% with no false alarms. The sensors withstood strong winds and rainy weather.