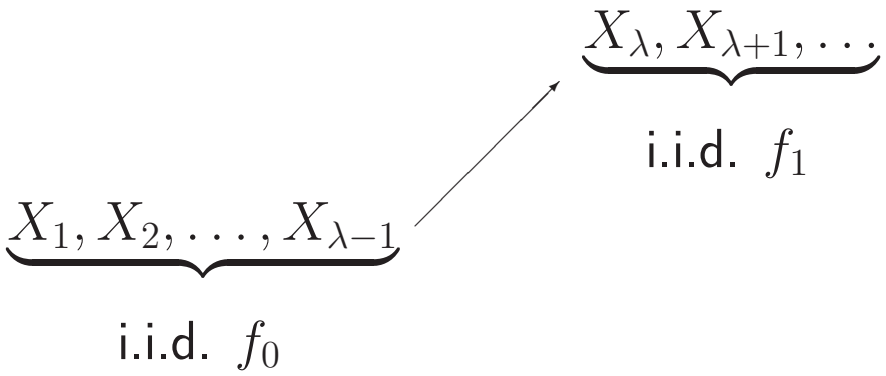
Asymptotic Analysis of Bayesian Quickest Change Detection Procedures

Venu Veeravalli
Dept of ECE and Coordinated Science Lab
University of Illinois at Urbana Champaign
vvv@uiuc.edu
http://www.comm.csl.uiuc.edu/∼vvv

Alexander Tartakovsky
Center for Applied Mathematical Sciences
University of Southern California, Los Angeles
tartakov@math.usc.edu

Michael Baron
University of Texas, Dallas and IBM Research
mbaron@utdallas.edu

Quickest Change Detection



- *Detection Procedure:* Stop at time τ and declare "change"
- *Applications:* failure detection, surveillance, process control,etc.
- *Lorden/Pollak Minimax Formulation*: Minimize worst case expected delay, subject to upper bound on false alarm rate
- *Shiryaev Bayesian Formulation*: Minimize expected delay subject to upper bound on false alarm probability
  - Assume prior distribution for λ – geometric

Minimax Formulation

- Supremum Detection Delay (SDD):

$$\mathrm{SDD} = \sup_{1 \leq \lambda < \infty} \mathrm{E}_\lambda \left[(\tau - \lambda) | \tau \geq \lambda \right]$$

- False Alarm Rate (FAR):

$$\mathrm{FAR} = 1/\mathrm{E}_\infty[\tau]$$

- *Optimization Problem:* Minimize SDD subject to $\mathrm{FAR} \leq \gamma$

- *Solution:* CUSUM test

$$\nu_a = \inf \left\{ n \geq 1 \ : \ \sum_{i=1}^{n} Y_i - \min_{k \leq n} \sum_{i=1}^{k} Y_i \ > \ a \right\}$$

  where

$$Y_i = \log[f_1(X_i)/f_0(X_i)]$$

- (Asymptotic) optimality [Lorden '71, Pollak '85, Moustakides '86]

Bayesian Formulation

- Prior

$$\pi(k) = \mathrm{P}\{\lambda = k\} = \rho(1-\rho)^{k-1}\,, \quad k \geq 1$$

- Average Detection Delay (ADD)

$$\mathrm{ADD}(\tau) = \mathrm{E}\left[(\tau - \lambda)|\tau \geq \lambda\right]$$

- Probability of False Alarm (PFA)

$$\mathrm{PFA}(\tau) = \mathrm{P}(\tau < \lambda)$$

- *Problem:* Minimize ADD subject to $\mathrm{PFA} \leq \alpha$

- Sufficient statistic

$$p_n = \mathrm{P}\left(\lambda \leq n \mid X_1, \ldots, X_n\right)$$

- *Shiryaev's Solution:*

$$\nu_A = \inf\{n \geqslant 1 : p_n \geqslant A\}$$

Relationship between CUSUM and Shiryaev Tests

- Define $\quad R_{\rho,n} := \dfrac{p_n}{\rho(1-p_n)}$

- Recursion

$$R_{\rho,n} = \frac{1}{1-\rho}[R_{\rho,n-1}+1]L_n, \quad R_{\rho,0} = 0$$

- *Shiryaev Test*

$$\nu_B = \inf\{n \geqslant 1 : R_{\rho,n} \geqslant B\}$$

- "Uniform" prior: Let $\rho \longrightarrow 0$, then

  - Shiryaev-Roberts Pollak (SRP) test statistic

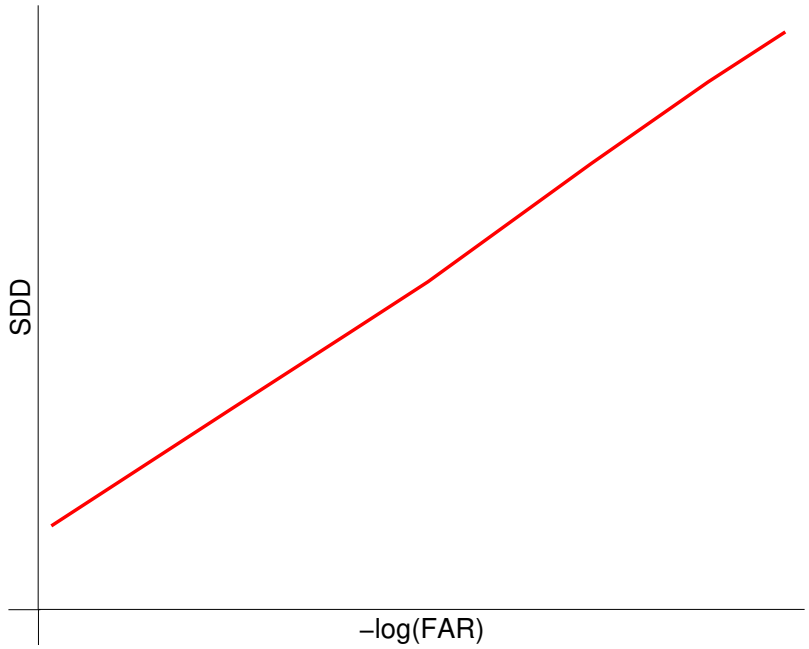
$$R_n = [R_{n-1}+1]L_n, \quad R_0 = 0$$

  - Performance $\approx$ CUSUM under minimax criterion [Pollak '85]

Performance Analysis of Minimax Procedures

- Asymptotic performance as FAR constraint $\gamma \to 0$.
- For CUSUM and SRP tests, if threshold such that $\mathrm{FAR} \approx \gamma$

$$\mathrm{SDD} \sim \frac{-\log \gamma}{D(f_1, f_0)}, \quad \text{as } \gamma \to 0$$



- Higher order asymptotics also available

Performance of Bayesian Procedures

- *Shiryaev Test*

$$\nu_B = \inf\{n \geqslant 1 : R_{\rho,n} \geqslant B\}$$

with

$$R_{\rho,n} = \frac{1}{1-\rho}[R_{\rho,n-1} + 1]L_n, \quad R_{\rho,0} = 0$$

- Define $\Delta(\alpha) := \{\tau : \mathrm{PFA}(\tau) \leq \alpha\}$

☞ *Bound on* PFA*:*

$$\mathrm{PFA}(\nu_B) \leq \frac{1}{1+\rho B}$$

☞ If $B = B_\alpha = \frac{1-\alpha}{\alpha\rho}$, then

$$\mathrm{PFA}(\nu_B) \leqslant \alpha \text{ , i.e., } \nu_B \in \Delta(\alpha)$$

First Order Asymptotics for ADD as $\alpha \to 0$

☞ *Result:* As $B \to \infty$

$$\mathrm{ADD}(\nu_B) \sim \frac{\log B}{D(f_1, f_0) + |\log(1-\rho)|}$$

- lower bound using Chebyshev's inequality
- upper bound by constructing simple test with worse asymptotic ADD

- Note: ν_B is optimum only if B is such that $\mathrm{PFA}(\nu_B) = \alpha$.

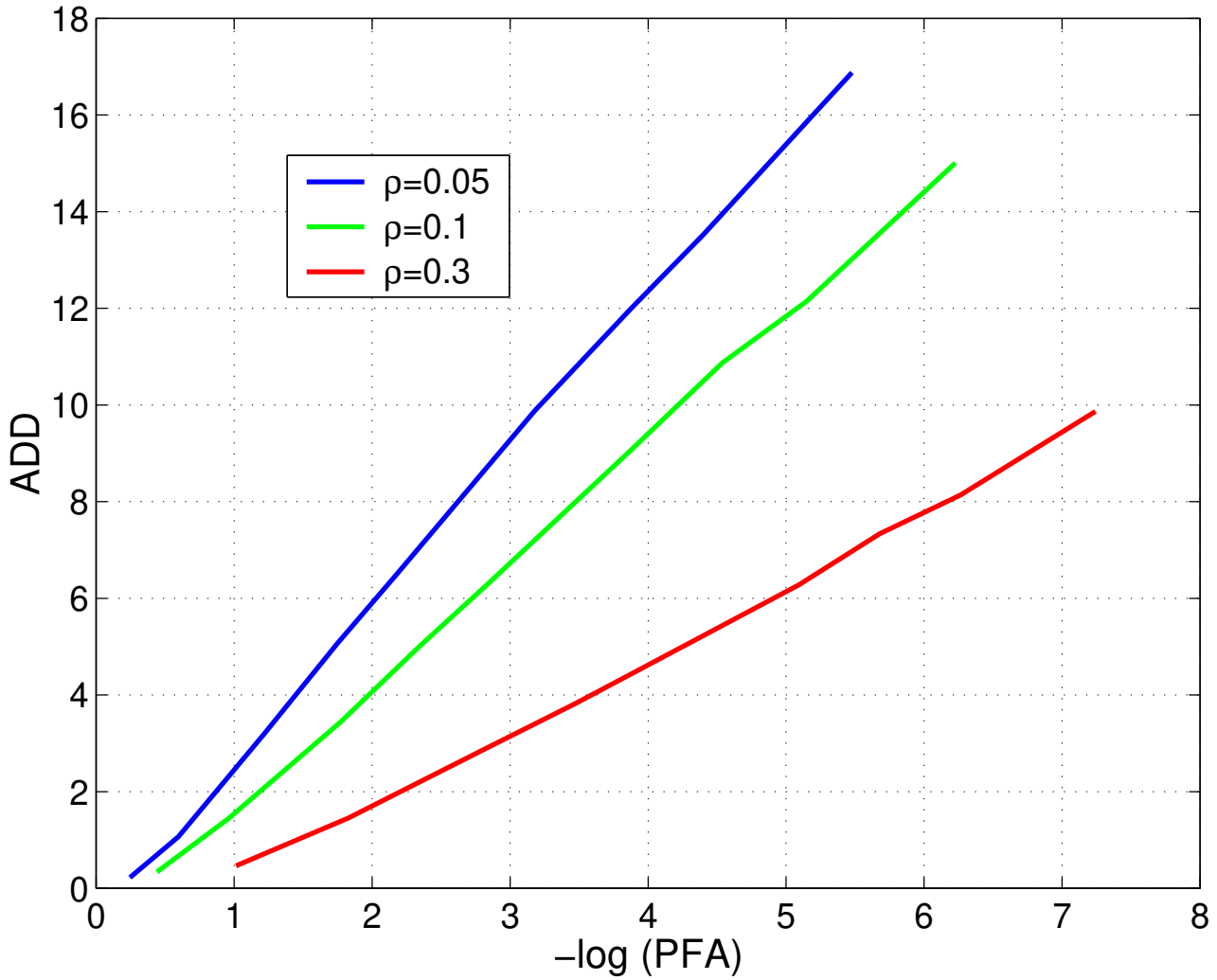
- What if $B = B_\alpha = \frac{1-\alpha}{\alpha\rho}$?

☞ *Asymptotic Optimality of* ν_{B_α} *:* As $\alpha \to 0$,

$$\inf_{\tau \in \mathbf{\Delta}(\alpha)} \mathrm{ADD}(\tau) \sim \mathrm{ADD}(\nu_{B_\alpha}) \sim \frac{-\log \alpha}{D(f_1, f_0) + |\log(1-\rho)|}$$

Simulation Results

Observations i.i.d. $\mathcal{N}(0,1)$ before and $\mathcal{N}(1,1)$ after change



Extension to non-iid Case

- Under measure P_k,

  pdf of X_n given X^{n-1} is $f_{0,n}(X_n|X^{n-1})$for $n \leq k-1$

  pdf of X_n given X^{n-1} is $f_{1,n}(X_n|X^{n-1})$for $n \geqslant k$

- Shiryaev test no longer optimal

- Log-likelihood ratio between P_k and P_∞ at time n

$$Z_n^k = \sum_{t=k}^{n} \log \frac{f_{1,t}(X_t|X^{t-1})}{f_{0,t}(X_t|X^{t-1})}$$

- *Assumption*

$$\frac{1}{n-k} Z_n^k \xrightarrow[n\to\infty]{\mathrm{P}_k-\text{completely}} q \quad \forall k$$

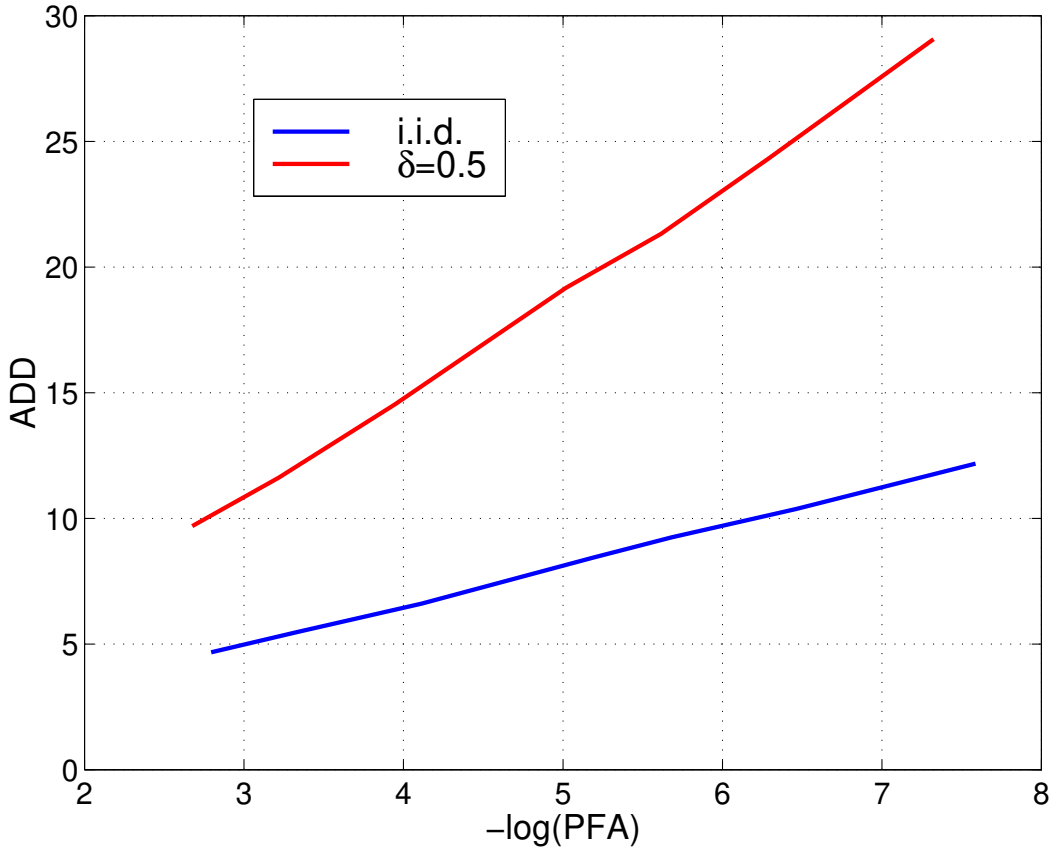
☞ *Asymptotic Optimality and Performance:* If $B = B_\alpha$, then as $\alpha \to 0$

$$\inf_{\tau\in\Delta(\alpha)} \mathrm{ADD}(\tau) \sim \mathrm{ADD}(\nu_B) \sim \frac{-\log\alpha}{q + |\log(1-\rho)|}$$

Example: Gauss-Markov Observations

- $X_n \sim \mathcal{N}(0, \sigma^2)$ before change, $X_n \sim \mathcal{N}(\theta, \sigma^2)$ after change, with corr. coeff. δ

$$\frac{1}{n-k} Z_n^k \xrightarrow[n\to\infty]{\mathrm{P}_k-\text{completely}} \frac{(1-\delta)^2\theta^2}{2\sigma^2} \quad \forall k$$



Further Results

- First order asymptotic optimality and performance extend to *conditional ADD*:

$$\mathrm{CADD}_k(\tau) = \mathrm{E}_k\left[(\tau - k)|\{\tau \geq k\}\right], \quad k = 1, 2, \ldots$$

- First order asymptotic optimality and performance extend to *higher order moments* of detection delay

- Accurate expressions for $\mathrm{PFA}(\nu_B)$ and $\mathrm{CADD}_k(\nu_B)$ in the iid case using *nonlinear renewal theory*. As $B \to \infty$

$$\mathrm{PFA}(\nu_B) = \frac{\zeta(\rho, D)}{\rho B}(1 + o(1))$$

$$\mathrm{CADD}_k(\nu_B) = \frac{\log B - C_k(\rho, D) + \overline{\varkappa}(\rho, D)}{D(f_1, f_0) + |\log(1 - \rho)|} - 1 + o(1)$$

Example: Higher Order Approximations

Observations i.i.d. $\mathcal{N}(0,1)$ before and $\mathcal{N}(1,1)$ after change, $\rho = 0.01$

$$\text{Set } B = \frac{\zeta(\rho, D)}{\rho\ \alpha}$$

α	PFA (sim)	CADD_1 (sim)	First Order	Higher Order
0.1000	0.0931	9.9694	11.3807	9.9577
0.0600	0.0589	11.0168	12.3823	10.9685
0.0300	0.0302	12.2502	13.7412	12.3090
0.0100	0.0098	14.4419	15.8952	14.3778
0.0060	0.0063	15.4933	16.8967	15.4721
0.0030	0.0031	16.8163	18.2557	16.7995
0.0010	0.0010	18.9732	20.4096	18.9628

Asymptotic Performance of CUSUM and SRP Tests

- CUSUM test:

$$\tau_a = \inf\left\{ n \geq 1 \ : \ \sum_{i=1}^{n} \log L_i - \min_{k \leq n} \sum_{i=1}^{k} \log L_i > a \right\}$$

- SRP test:

$$\tau_b = \inf\{ n \geqslant 1 : q_n \geqslant b \}$$

with

$$q_k = [q_{k-1} + 1] L_k, \quad q_0 = 0$$

- Can set a, b s.t. $\mathrm{PFA}(\nu_a) \approx \alpha$ and $\mathrm{PFA}(\nu_b) \approx \alpha$

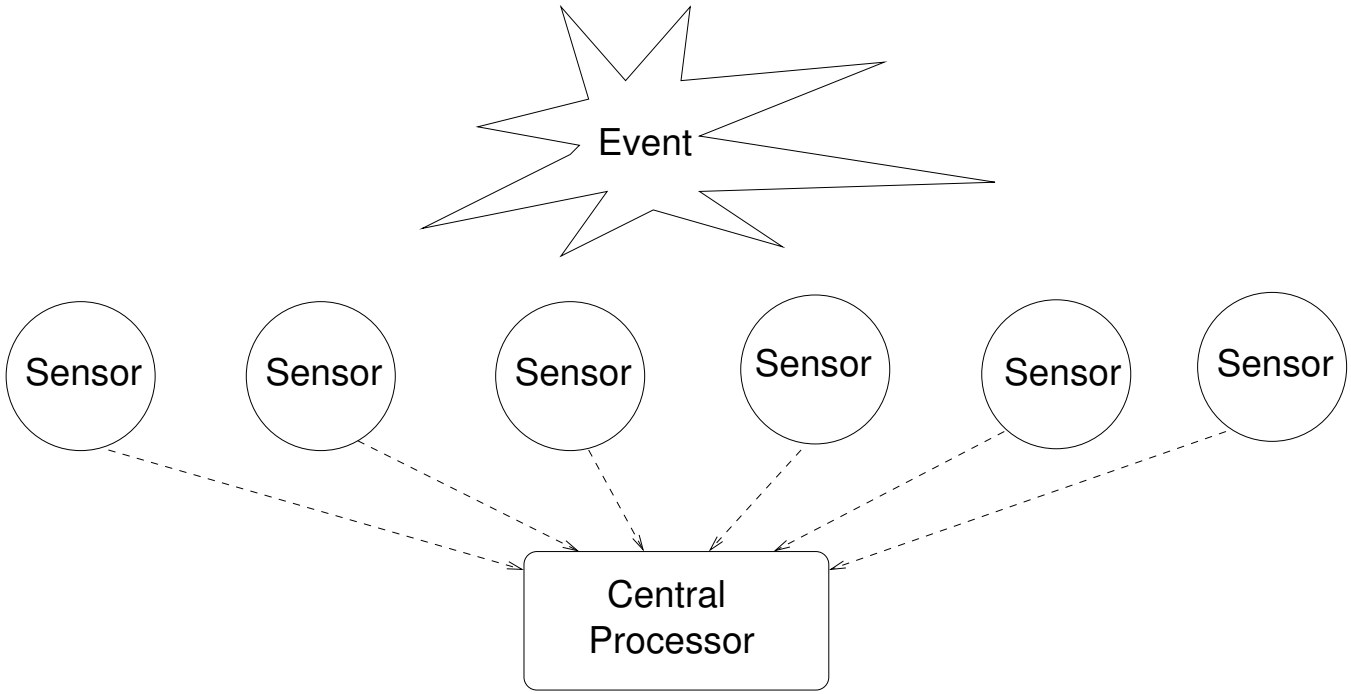
- For both tests, as $\alpha \to 0$,

$$\mathrm{ADD} = \frac{-\log \alpha}{D(f_1, f_0)} (1 + o(1))$$

Thus

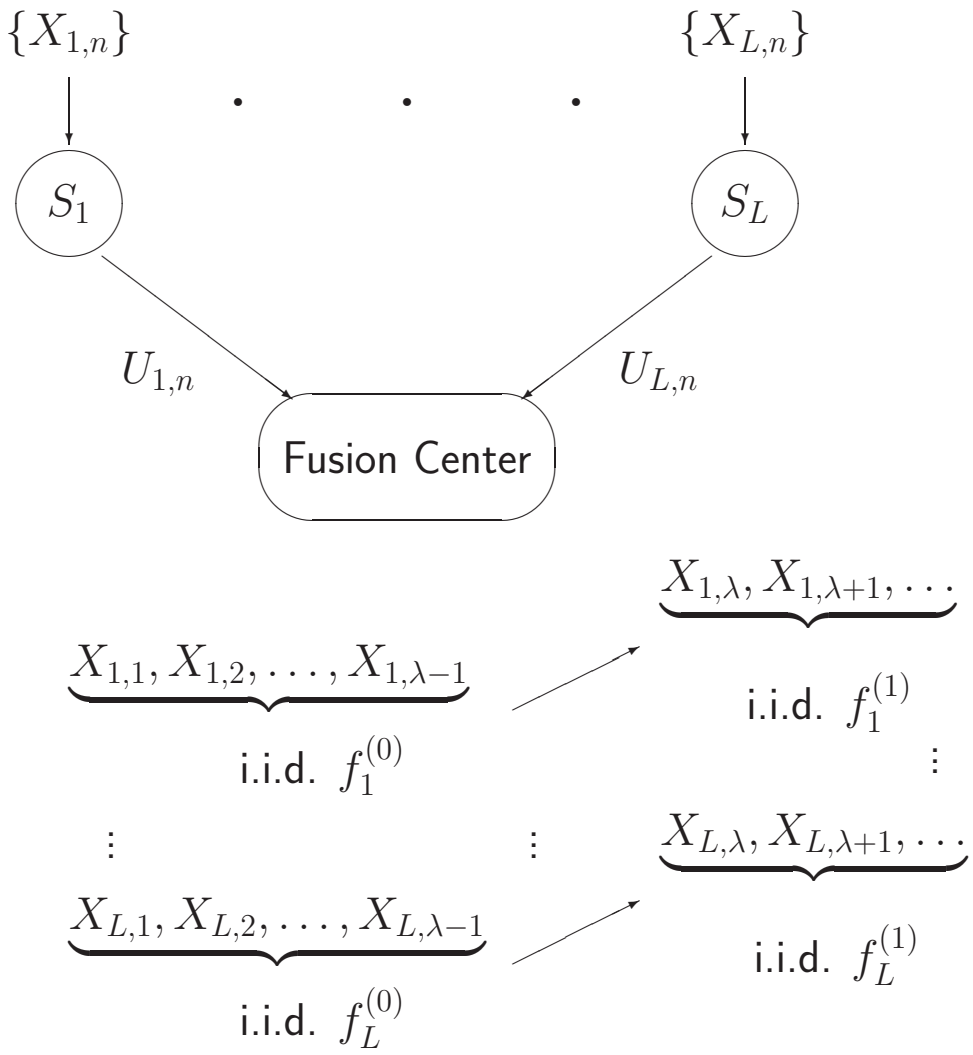
$$\lim_{\alpha \to 0} \frac{\mathrm{ADD}(\textsf{CUSUM}, \textsf{SRP})}{\mathrm{ADD}(\textsf{Shiryaev})} = 1 + \frac{|\log(1-\rho)|}{D(f_1, f_0)}$$

Change Detection Using Distributed Sensors



- Low-cost, limited storage, wireless sensors
- *Applications:* environmental awareness, surveillance, security

Decentralized Change Detection



$$U_{\ell,n} = \psi_{\ell,n}(X_{\ell,n})$$

- $\star$ $U_{\ell,n} \in$ finite alphabet (comm. constraints)
- $\star$ Need to jointly optimize:
  - Fusion center τ
  - Sensor quantizers $\{\psi_{\ell,n}\}$

Optimization Problem

- *Detection policy:* $\phi = \left(\tau,\ \psi_n^{(\ell)}, \ell = 1, \ldots, L, k = 1, 2, \ldots\right)$
- *Goal:* Choose ϕ to minimize ADD subject to $\mathrm{PFA} \leq \alpha$
- Optimum solution intractable!
- $g_{\ell,n}^{(j)}$ denotes the pmf induced on $U_{\ell,n}$ when $X_{\ell,n}$ had pdf $f_\ell^{(j)}$
- Log-likelihood ratio between P_k and P_∞ and sensor ℓ at time n

$$Z_n^k(\ell) = \sum_{i=k}^{n} \log \frac{g_{\ell,i}^{(1)}(U_{\ell,i})}{g_{\ell,i}^{(0)}(U_{\ell,i})}$$

- Restrict optimization to sensor quantizers for which

$$\frac{1}{n-k}\, Z_n^k(\ell) \xrightarrow[n\to\infty]{\mathrm{P}_k-\text{completely}} q_\ell$$

$$q_\ell \leq \max_{\psi_{\ell,n}} D(g_{\ell,n}^{(1)}, g_{\ell,n}^{(0)})$$

Asymptotically Optimum Solution

- *Result:* For fixed sensor quantizers, Shiryaev's test ν_B on sensor outputs is asymptotically optimum

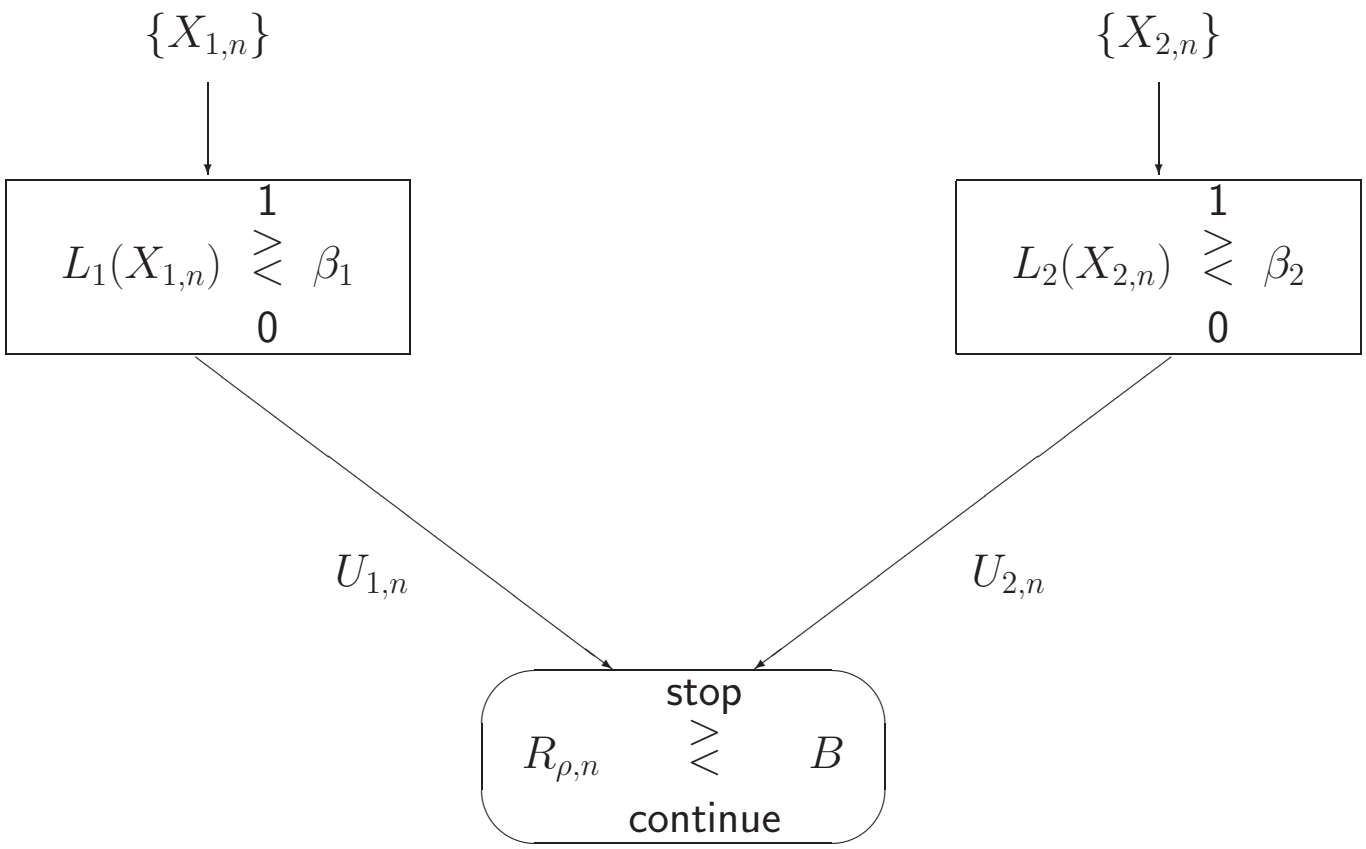
$$\inf_{\tau \in \Delta(\alpha)} \mathsf{ADD}(\tau) \sim \mathsf{ADD}(\nu_{B_\alpha}) \sim \frac{-\log \alpha}{\sum_{\ell=1}^{L} q_\ell + |\log(1-\rho)|} \quad \text{as } \alpha \to 0$$

- *Corollary:* It is asymptotically optimum (as $\alpha \to 0$) for sensor S_ℓ at time n to pick $\psi_{\ell,n}$ to maximize $D(g_{\ell,n}^{(1)}, g_{\ell,n}^{(0)})$

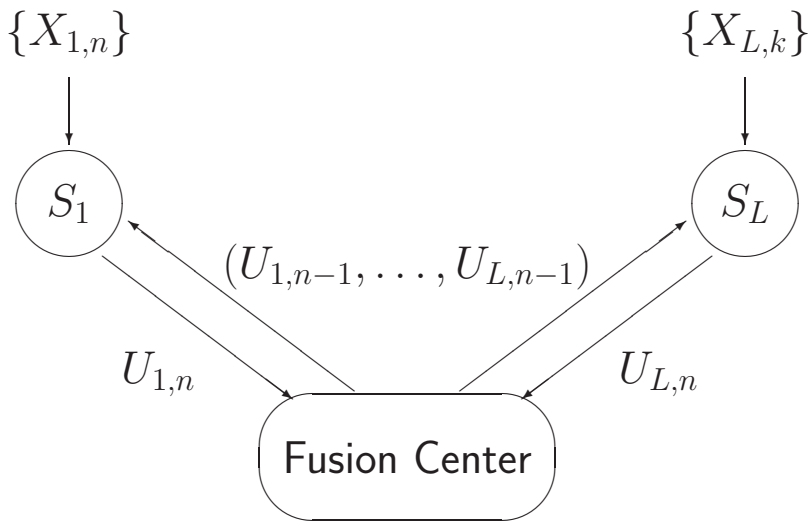
- Solution to $D(g_{\ell,n}^{(1)}, g_{\ell,n}^{(0)})$ maximization gives optimum $\psi_{\ell,n}$ to be monotone likelihood ratio quantizer (MLRQ) that is *stationary* in n, say ψ_ℓ^{opt}

- *Result:* Asymptotically optimum policy ϕ_{opt} for decentralized change detection consists of stationary (in time) set of MLRQ's at sensors, followed by Shiryaev's procedure based on $\{U_{1,n}, \ldots, U_{L,n}\}_{n \geqslant 1}$ at fusion center

Asymptotically Optimum Policy Structure

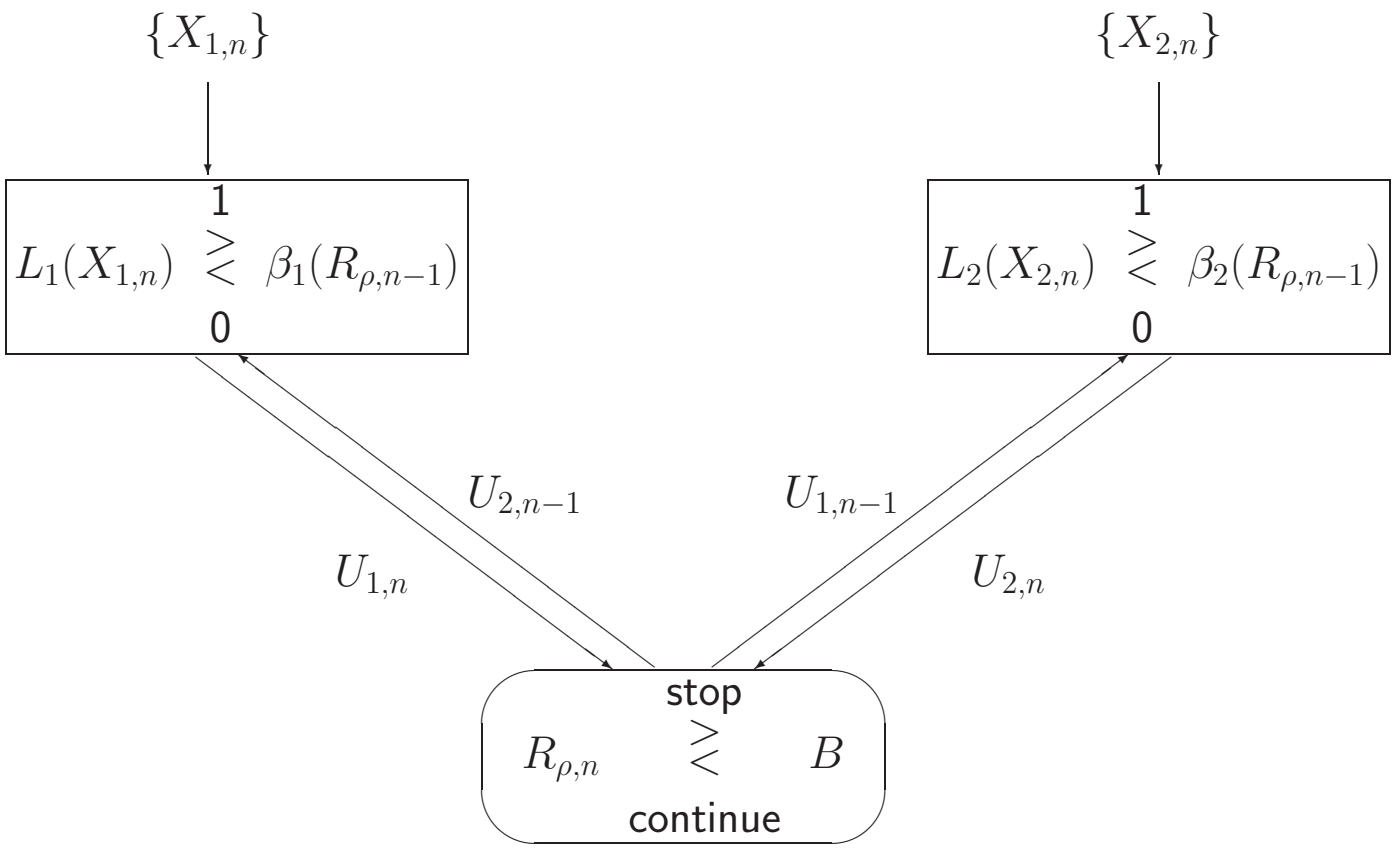


Decision Feedback From Fusion Center



- Introduce feedback to aid tractability
- Optimum solution can be found via dynamic programming [VVV, IT Nov '01]
- Optimum performance with feedback upper bound on performance without feedback

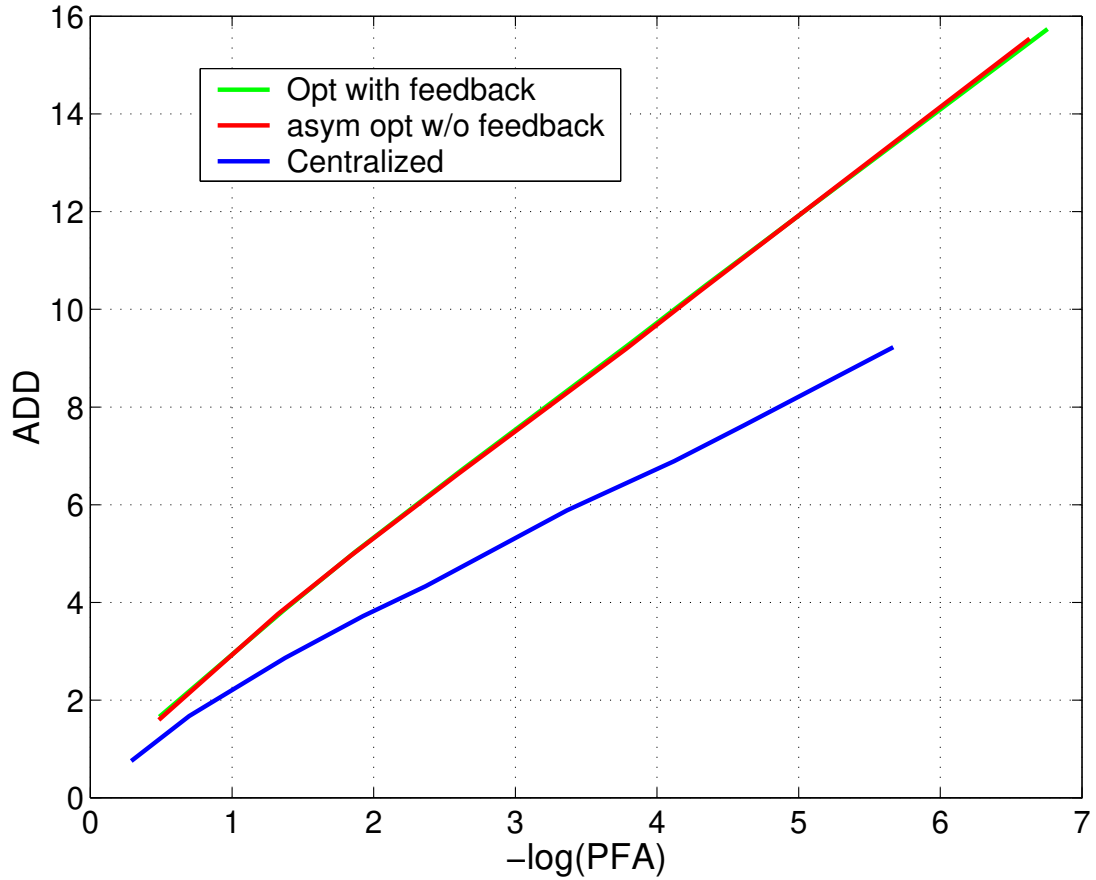
Structure of Optimum Policy with Feedback



$$R_{\rho,n} = \frac{1}{1-\rho}\,[R_{\rho,n-1}+1]\prod_{\ell=1}^{2} L_\ell^U(U_{\ell,n}, R_{\rho,n-1}), \quad \text{with } R_{\rho,0}=0$$

Decision Feedback – Numerical Results

Sensor observations $\mathcal{N}(0,1)$ before and $\mathcal{N}(0.5,1)$ after change; $\rho = 0.05, L = 5$



☞ *Result:* Can show that as $\alpha \to 0$, optimum solution with feedback converges to solution without feedback

Conclusions

- Asymptotic analysis of Bayesian change point detection procedures
- Analysis for iid as well non-iid cases
- Higher order approximations for PFA and ADD in i.i.d case
- Asymptotically optimum solution to decentralized change detection problem without feedback
- As $\mathrm{PFA} \rightarrow 0$, feedback from fusion center does not improve performance in decentralized change detection