Spectrum Sensing in Cognitive Radios using Distributed Sequential Detection

Jithin K. S.

Research Supervisor: Prof. Vinod Sharma
External Examiner: Prof. Shankar Prakriya, IIT Delhi

ECE Dept., Indian Institute of Science, Bangalore

February 4, 2013

Outline

- Introduction
- Parametric Distributed Sequential Tests
- Nonparametric Distributed Sequential Tests
- Multihypothesis scenario
- Conclusions

Introduction

Introduction

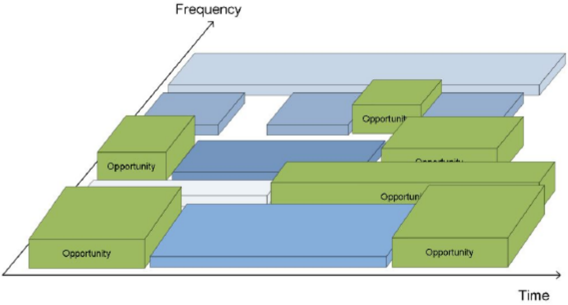
- CR: Radio systems, which perform Radio Environment Analysis, identify spectral holes and operate in those holes.

Introduction

- CR: Radio systems, which perform Radio Environment Analysis, identify spectral holes and operate in those holes.
- Licensed (Primary) and Unlicensed (Secondary) users

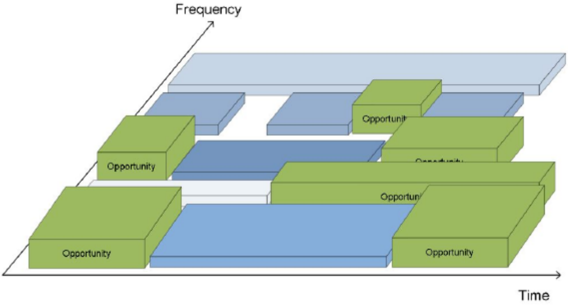
Introduction

- CR: Radio systems, which perform Radio Environment Analysis, identify spectral holes and operate in those holes.
- Licensed (Primary) and Unlicensed (Secondary) users
- Spectral holes in various dimensions



Introduction

- CR: Radio systems, which perform Radio Environment Analysis, identify spectral holes and operate in those holes.
- Licensed (Primary) and Unlicensed (Secondary) users
- Spectral holes in various dimensions



- Aided by evolving SDR technology

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:
  - Matched filter

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:
  - Matched filter
  - Cyclostationary detector

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:
  - Matched filter
  - Cyclostationary detector
  - Energy detector

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:
  - Matched filter
  - Cyclostationary detector
  - Energy detector
- Challenges:

Spectrum Sensing

- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:
  - Matched filter
  - Cyclostationary detector
  - Energy detector
- Challenges:
  - Shadowing, Fading and low SNR

Spectrum Sensing

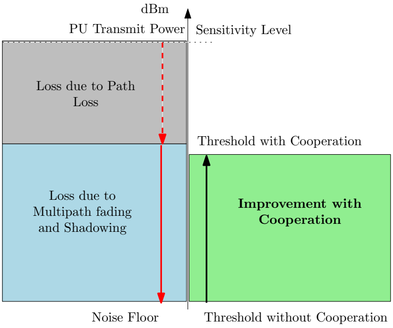
- Identify spectral holes and quickly detect the onset of primary transmission
- Spectrum Sensing Hypothesis testing formulation:
  H_0: **Primary user transmission is absent**
  H_1: **Primary user transmission is present**
- Primary Techniques:
  - Matched filter
  - Cyclostationary detector
  - Energy detector
- Challenges:
  - Shadowing, Fading and low SNR
  - Sensing Frequency and duration

Cooperative setup

- Mitigates multipath fading, shadowing and hidden node problem.
- Improves Probability of Errors (P_{FA} and P_{MD}) and Expected Detection Delay (E_{DD})
- A distributed statistical inference problem

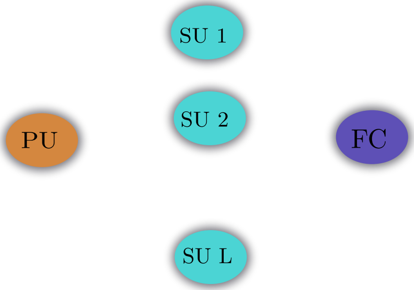
Cooperative setup

- Mitigates multipath fading, shadowing and hidden node problem.
- Improves Probability of Errors (P_{FA} and P_{MD}) and Expected Detection Delay (E_{DD})
- A distributed statistical inference problem



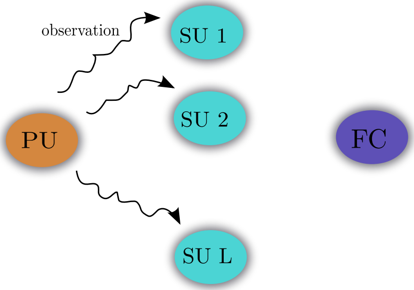
Problems addressed in this thesis

- Distributed Detection:



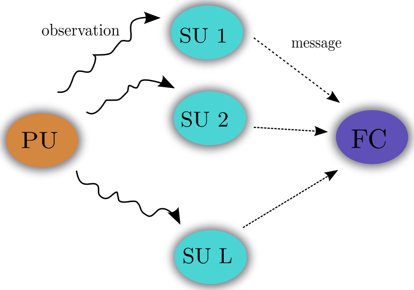
Problems addressed in this thesis

- Distributed Detection:



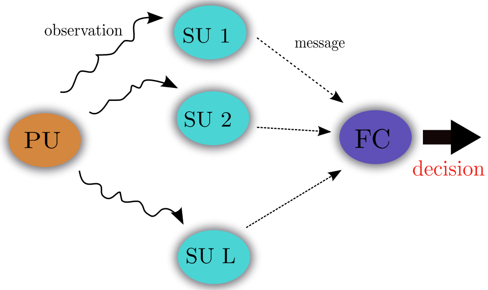
Problems addressed in this thesis

- Distributed Detection:



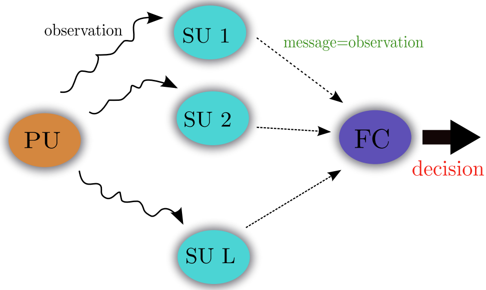
Problems addressed in this thesis

- Distributed Detection:



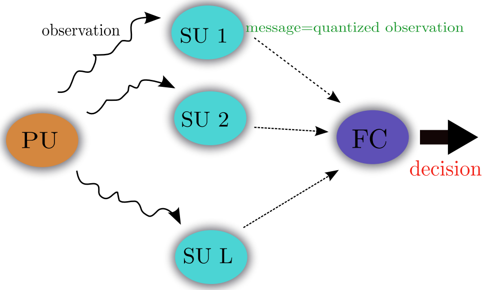
Problems addressed in this thesis

- Distributed Detection: centralized



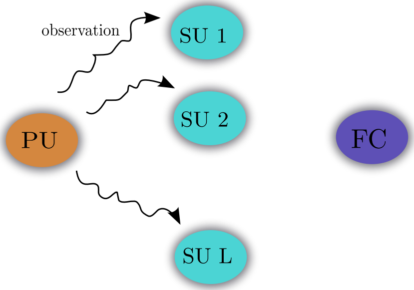
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized



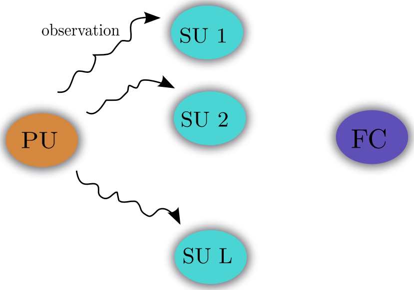
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR:



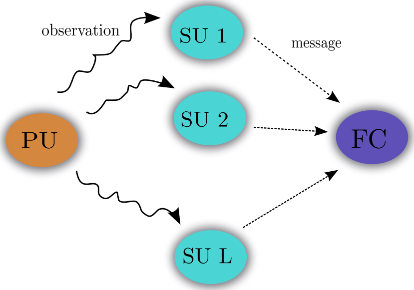
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: synchronous



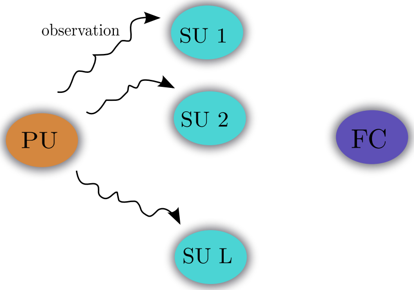
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: synchronous



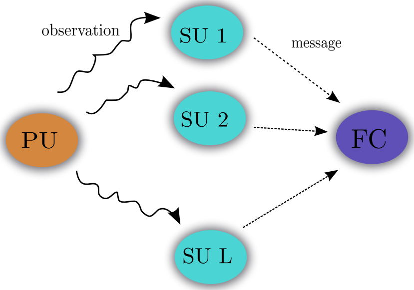
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: synchronous



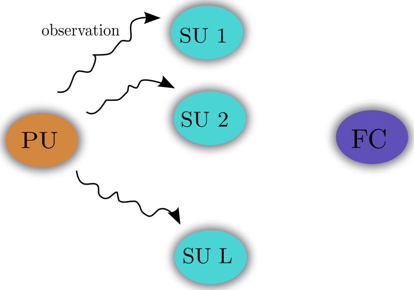
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: synchronous



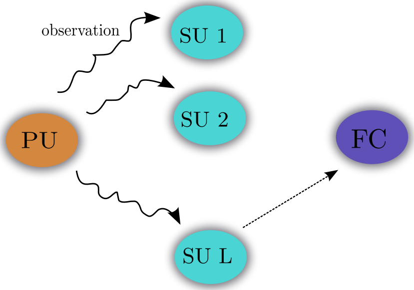
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: ~~synchronous~~ /Asynchronous



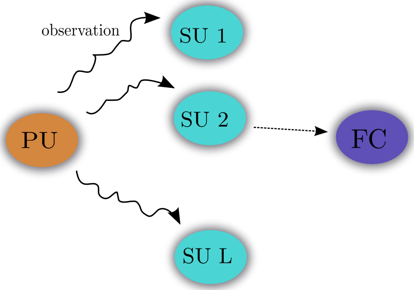
Problems addressed in this thesis

- Distributed Detection:
  ~~centralized~~/decentralized
- Sampling at CR:
  ~~synchronous~~ /Asynchronous



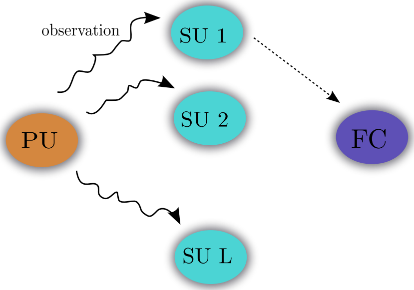
Problems addressed in this thesis

- Distributed Detection:
  ~~centralized~~/decentralized
- Sampling at CR:
  ~~synchronous~~ /Asynchronous



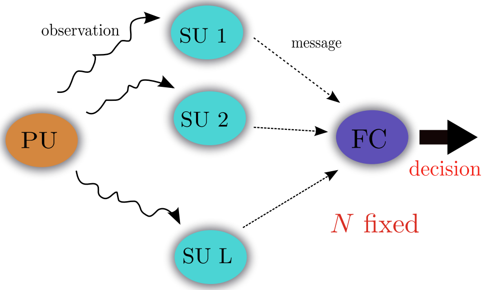
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: ~~synchronous~~ /Asynchronous



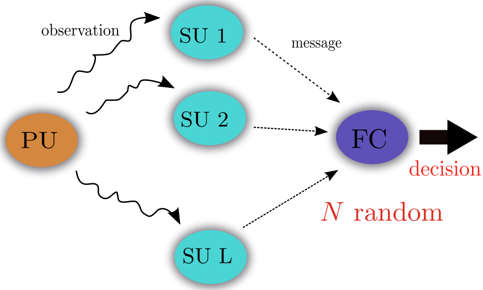
Problems addressed in this thesis

- Distributed Detection: ~~centralized~~/decentralized
- Sampling at CR: ~~synchronous~~ /Asynchronous
- Decision Mechanism: fixed sample size



Problems addressed in this thesis

- Distributed Detection:
  ~~centralized~~/decentralized
- Sampling at CR:
  ~~synchronous~~ /Asynchronous
- Decision Mechanism:
  ~~fixed sample size~~ /Sequential



Problems addressed in this thesis (Contd.)

- Focus on reducing the number of samples in the cooperative setup with reliability constraints
  (Decentralized Sequential Hypothesis Testing)

Problems addressed in this thesis (Contd.)

- Focus on reducing the number of samples in the cooperative setup with reliability constraints
  (Decentralized Sequential Hypothesis Testing)
  - Fast detection of idle TV bands

Problems addressed in this thesis (Contd.)

- Focus on reducing the number of samples in the cooperative setup with reliability constraints
  (Decentralized Sequential Hypothesis Testing)
  - Fast detection of idle TV bands
  - Quickest detection of unoccupied spectrum in slotted primary user transmission systems

Problems addressed in this thesis (Contd.)

- Focus on reducing the number of samples in the cooperative setup with reliability constraints
  (Decentralized Sequential Hypothesis Testing)
  - Fast detection of idle TV bands
  - Quickest detection of unoccupied spectrum in slotted primary user transmission systems
- Imprecise estimates of the channel gains and noise power
  (parameter uncertainty-Composite Sequential Testing)

Problems addressed in this thesis (Contd.)

- Focus on reducing the number of samples in the cooperative setup with reliability constraints
  (Decentralized Sequential Hypothesis Testing)
  - Fast detection of idle TV bands
  - Quickest detection of unoccupied spectrum in slotted primary user transmission systems
- Imprecise estimates of the channel gains and noise power
  (parameter uncertainty-Composite Sequential Testing)
- Channel gain statistic unknown (Universal Sequential Testing)

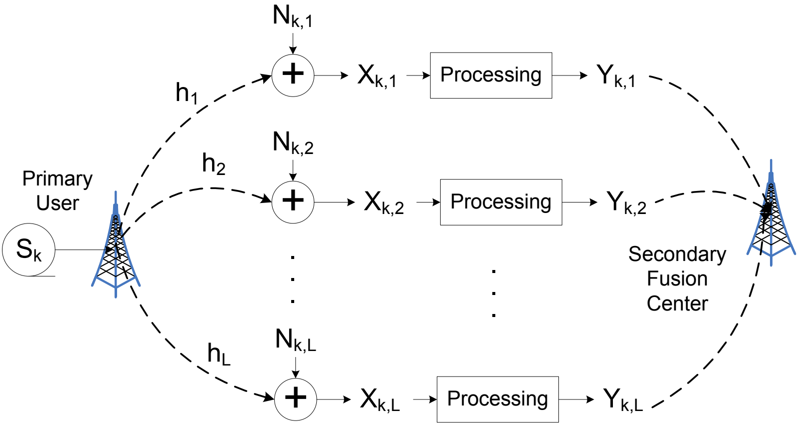
Problems addressed in this thesis (Contd.)

- Focus on reducing the number of samples in the cooperative setup with reliability constraints
  (Decentralized Sequential Hypothesis Testing)
  - Fast detection of idle TV bands
  - Quickest detection of unoccupied spectrum in slotted primary user transmission systems
- Imprecise estimates of the channel gains and noise power
  (parameter uncertainty-Composite Sequential Testing)
- Channel gain statistic unknown (Universal Sequential Testing)
- Detect the spectrum and identify the primary user
  (Decentralized Multihypothesis Sequential Testing)

Related Work

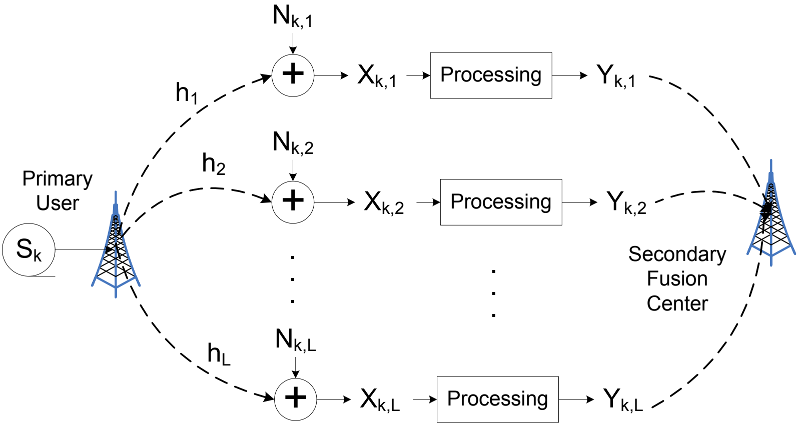
- Sequential tests: Wald'47, Irle'84
- Decentralized Sequential tests: Mei'08, Fellouris et al.'12, Yilmaz et al.'12
- Cooperative Spectrum Sensing and Censored detection: Akyildiz et al.'11
- Universal hypothesis tests: Levitan et al.'02, Jacob et al.'08, Unnikrishnan et al.'11
- Multihypothesis Sequential tests: Draglia et al.'99, Tartakovsky'00

Model



$N_{k,1}$
h_1
$X_{k,1}$
Processing
$Y_{k,1}$
$N_{k,2}$
h_2
$X_{k,2}$
Processing
$Y_{k,2}$
Primary User
S_k
Secondary Fusion Center
h_L
$N_{k,L}$
$X_{k,L}$
Processing
$Y_{k,L}$

Model



General Problem: N = Stopping time, δ_N = decision rule

$$\min_{N,\delta_N} E[N|H_0] \text{ and } \min_{N,\delta_N} E[N|H_1],$$

subject to $P_{FA} \le \alpha_1$ and $P_{MD} \le \alpha_2$

Parametric Distributed Sequential Tests

- At l $^{\text{th}}$ CR,

$$\begin{aligned} H_0 : \quad & X_{k,l} \sim f_{0,l} \\ H_1 : \quad & X_{k,l} \sim f_{1,l} \end{aligned}$$

- $f_{0,l}$ and $f_{1,l}$ are fully known
- $f_{0,l}$ and $f_{1,l}$ are not fully known: parametric family.

DualSPRT Algorithm

1. At Node, l, Sequential Probability Ratio Test (SPRT),

$$W_{k,l} = W_{k-1,l} + \log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})},\ W_{0,l} = 0.$$

DualSPRT Algorithm

1. At Node, l, Sequential Probability Ratio Test (SPRT),

$$W_{k,l} = W_{k-1,l} + \log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})},\ W_{0,l} = 0.$$

2. Node l transmits

$$Y_{k,l} = b_1 1_{\{W_{k,l} \geq \gamma_1\}} + b_0 1_{\{W_{k,l} \leq -\gamma_0\}}$$

DualSPRT Algorithm

1. At Node, l, Sequential Probability Ratio Test (SPRT),

$$W_{k,l} = W_{k-1,l} + \log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})}, \; W_{0,l} = 0.$$

2. Node l transmits

$$Y_{k,l} = b_1 1_{\{W_{k,l} \geq \gamma_1\}} + b_0 1_{\{W_{k,l} \leq -\gamma_0\}}$$

Noisy MAC at the FC, Z_k i.i.d. MAC noise

$$Y_k = \sum_{l=1}^{L} Y_{k,l} + Z_k$$

DualSPRT Algorithm

1. At Node, l, Sequential Probability Ratio Test (SPRT),

$$W_{k,l} = W_{k-1,l} + \log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})}, \; W_{0,l} = 0.$$

2. Node l transmits

$$Y_{k,l} = b_1 1_{\{W_{k,l} \geq \gamma_1\}} + b_0 1_{\{W_{k,l} \leq -\gamma_0\}}$$

   Noisy MAC at the FC, Z_k i.i.d. MAC noise

$$Y_k = \sum_{l=1}^{L} Y_{k,l} + Z_k$$

3. FC runs SPRT: g_μ is p.d.f. of $\mu + Z_k$

DualSPRT Algorithm

1. At Node, l, Sequential Probability Ratio Test (SPRT),

$$W_{k,l} = W_{k-1,l} + \log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})}, \; W_{0,l} = 0.$$

2. Node l transmits

$$Y_{k,l} = b_1 1_{\{W_{k,l} \geq \gamma_1\}} + b_0 1_{\{W_{k,l} \leq -\gamma_0\}}$$

   Noisy MAC at the FC, Z_k i.i.d. MAC noise

$$Y_k = \sum_{l=1}^{L} Y_{k,l} + Z_k$$

3. FC runs SPRT: g_μ is p.d.f. of $\mu + Z_k$

$$F_k = F_{k-1} + \log \frac{g_{\mu_1}(Y_k)}{g_{-\mu_0}(Y_k)}, \; F_0 = 0,$$

DualSPRT Algorithm

1. At Node, l, Sequential Probability Ratio Test (SPRT),

$$W_{k,l} = W_{k-1,l} + \log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})}, \; W_{0,l} = 0.$$

2. Node l transmits

$$Y_{k,l} = b_1 1_{\{W_{k,l} \geq \gamma_1\}} + b_0 1_{\{W_{k,l} \leq -\gamma_0\}}$$

   Noisy MAC at the FC, Z_k i.i.d. MAC noise

$$Y_k = \sum_{l=1}^{L} Y_{k,l} + Z_k$$

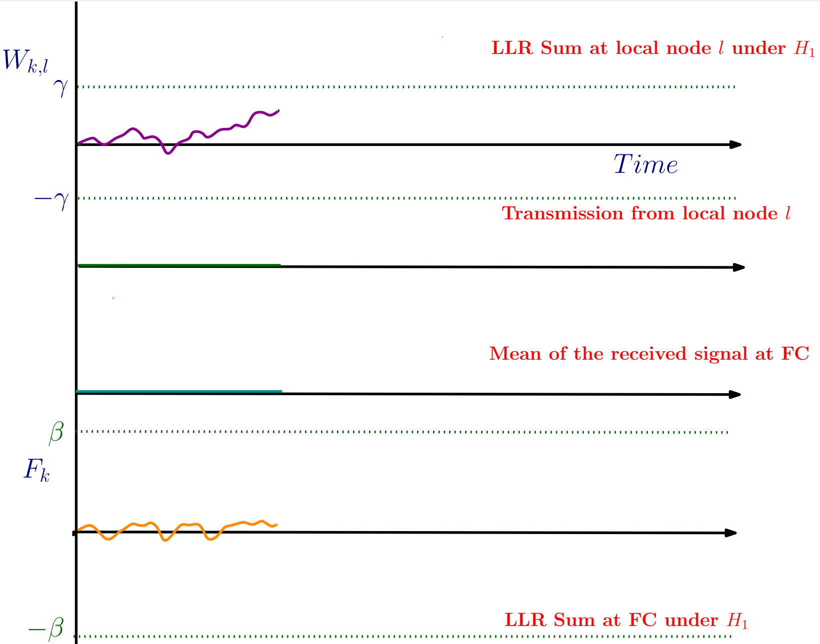
3. FC runs SPRT: g_μ is p.d.f. of $\mu + Z_k$

$$F_k = F_{k-1} + \log \frac{g_{\mu_1}(Y_k)}{g_{-\mu_0}(Y_k)}, \; F_0 = 0,$$

4. FC decides, at $N = \inf\{k : F_k \notin (-\beta_0, \beta_1)\}$, H_1 if $F_N \geq \beta_1$, H_0 if $F_N \geq -\beta_0$.

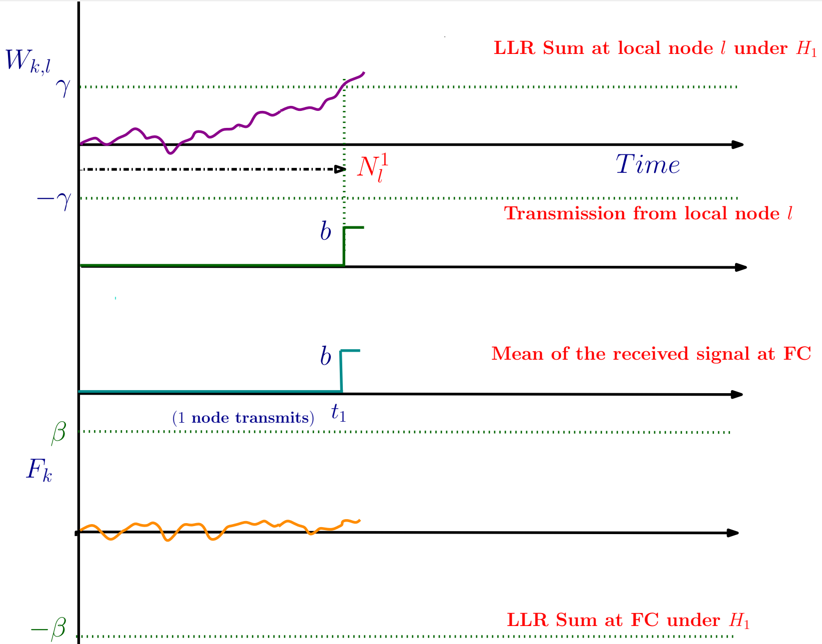
Performance Analysis of DualSPRT- E_{DD}

Sample path argument



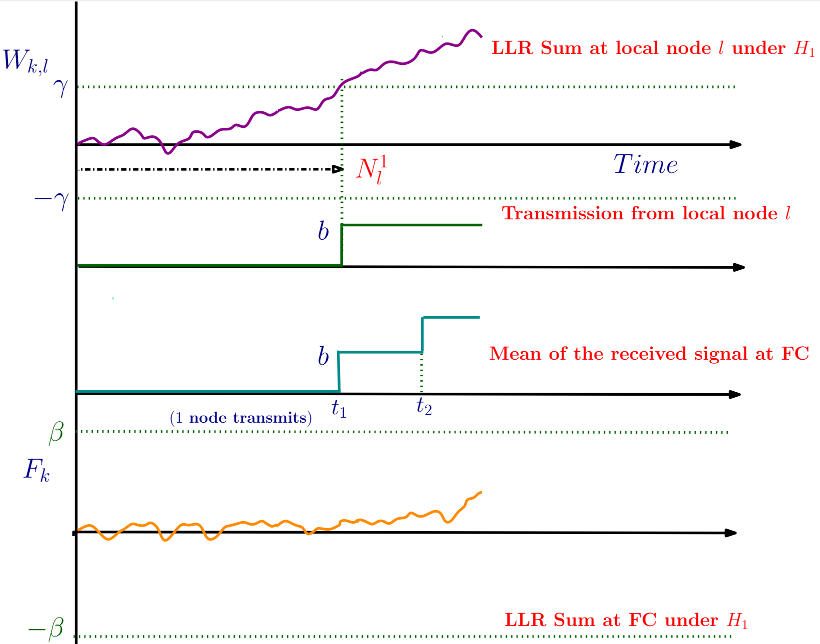
Performance Analysis of DualSPRT- E_{DD}

Sample path argument



Performance Analysis of DualSPRT- E_{DD}

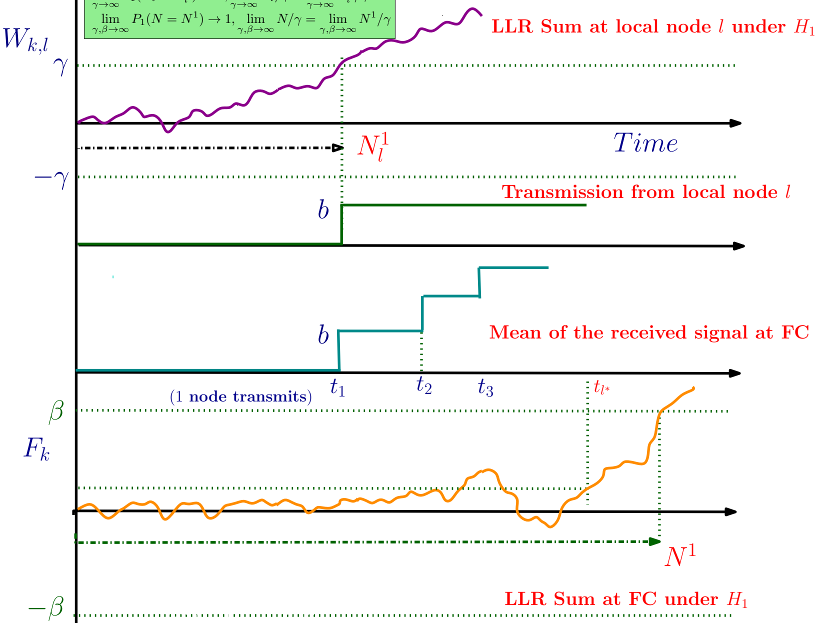
Sample path argument



Performance Analysis of DualSPRT- E_{DD}

Sample path argument

$$\lim_{\gamma\to\infty} P_1(N_l = N_l^1) \to 1, \lim_{\gamma\to\infty} N_l/\gamma = \lim_{\gamma\to\infty} N_l^1/\gamma$$
$$\lim_{\gamma,\beta\to\infty} P_1(N = N^1) \to 1, \lim_{\gamma,\beta\to\infty} N/\gamma = \lim_{\gamma,\beta\to\infty} N^1/\gamma$$



Performance Analysis of DualSPRT- E_{DD} (Contd.)

- At each node l, time to cross the threshold

$$N_l^1 \sim \mathcal{N}(\frac{\gamma}{\delta_l}, \frac{\rho_l^2 \gamma}{\delta_l^3}),\ \delta_l = E_1[LLR_l] \text{ and } \rho_l^2 = Var_1[LLR_l]$$

Performance Analysis of DualSPRT- E_{DD} (Contd.)

- At each node l, time to cross the threshold

$$N_l^1 \sim \mathcal{N}\left(\frac{\gamma}{\delta_l}, \frac{\rho_l^2 \gamma}{\delta_l^3}\right), \ \delta_l = E_1[LLR_l] \text{ and } \rho_l^2 = Var_1[LLR_l]$$

$$t_j = j\text{th order statistics of } N_1^1, N_2^1, \ldots, N_L^1$$

Performance Analysis of DualSPRT- E_{DD} (Contd.)

- At each node l, time to cross the threshold

$$N_l^1 \sim \mathcal{N}(\frac{\gamma}{\delta_l}, \frac{\rho_l^2 \gamma}{\delta_l^3}), \; \delta_l = E_1[LLR_l] \text{ and } \rho_l^2 = Var_1[LLR_l]$$

$$t_j = j\text{th order statistics of } N_1^1, N_2^1, \ldots, N_L^1$$

- $l^* = min\{j : \delta_{FC}^j > 0 \text{ and } \frac{\beta - \bar{F}_j}{\delta_{FC}^j} < E[t_{j+1}] - E[t_j]\}$

Performance Analysis of DualSPRT- E_{DD} (Contd.)

- At each node l, time to cross the threshold

$$N_l^1 \sim \mathcal{N}(\frac{\gamma}{\delta_l}, \frac{\rho_l^2 \gamma}{\delta_l^3}),\ \delta_l = E_1[LLR_l] \text{ and } \rho_l^2 = Var_1[LLR_l]$$

$$t_j = j\text{th order statistics of } N_1^1, N_2^1, \ldots, N_L^1$$

- $l^* = min\{j : \delta_{FC}^j > 0 \text{ and } \frac{\beta - \bar{F}_j}{\delta_{FC}^j} < E[t_{j+1}] - E[t_j]\}$
- $\bar{F}_j = E[F_{t_j - 1}]$, $\bar{F}_j = \bar{F}_{j-1} + \delta_{FC}^j \left(E[t_j] - E[t_{j-1}]\right)$, $\bar{F}_0 = 0$.

$$E_{DD}\left(E_1[N]\right) \approx E_1[N^1] \approx E[t_{l^*}] + \frac{\beta - \bar{F}_{l^*}}{\delta_{FC}^{l^*}}$$

Performance Analysis of DualSPRT- P_{MD}

$$P_{MD} = P_1(N^0 < N^1) \geq P_1(N^0 < t_1, N^1 > t_1)$$

Performance Analysis of DualSPRT- P_{MD}

$$\begin{aligned} P_{MD} = P_1(N^0 < N^1) &\geq P_1(N^0 < t_1, N^1 > t_1) \\ &\approx P_1(N^0 < t_1) \end{aligned}$$

Performance Analysis of DualSPRT- P_{MD}

$$
\begin{aligned}
P_{MD} = P_1(N^0 < N^1) &\geq P_1(N^0 < t_1, N^1 > t_1) \\
&\approx P_1(N^0 < t_1) \\
\text{Also } P_1(N^0 < N^1) &\leq P_1(N^0 < \infty)
\end{aligned}
$$

Performance Analysis of DualSPRT- P_{MD}

$$
\begin{aligned}
P_{MD} = P_1(N^0 < N^1) &\geq P_1(N^0 < t_1, N^1 > t_1) \\
&\approx P_1(N^0 < t_1) \\
\text{Also } P_1(N^0 < N^1) &\leq P_1(N^0 < \infty) \\
&= P_1(N^0 < t_1) + P_1(t_1 \leq N^0 < t_2) \\
&\quad + P_1(t_2 \leq N^0 < t_3) + \dots
\end{aligned}
$$

Performance Analysis of DualSPRT- P_{MD}

$$\begin{aligned} P_{MD} = P_1(N^0 < N^1) &\geq P_1(N^0 < t_1, N^1 > t_1) \\ &\approx P_1(N^0 < t_1) \\ \text{Also } P_1(N^0 < N^1) &\leq P_1(N^0 < \infty) \\ &= P_1(N^0 < t_1) + P_1(t_1 \leq N^0 < t_2) \\ &\quad + P_1(t_2 \leq N^0 < t_3) + \ldots \end{aligned}$$

$$P_{MD} \approx P_1(N^0 < t_1)$$

Performance Analysis of DualSPRT- P_{MD}

$$
\begin{aligned}
P_{MD} = P_1(N^0 < N^1) &\geq P_1(N^0 < t_1, N^1 > t_1) \\
&\approx P_1(N^0 < t_1) \\
\text{Also } P_1(N^0 < N^1) &\leq P_1(N^0 < \infty) \\
&= P_1(N^0 < t_1) + P_1(t_1 \leq N^0 < t_2) \\
&\quad + P_1(t_2 \leq N^0 < t_3) + \dots
\end{aligned}
$$

$$
\begin{aligned}
P_{MD} &\approx {\color{red} P_1(N^0 < t_1)} \\
&= \sum_{k=1}^{\infty} P\left[\{F_k < -\beta\} \cap_{n=1}^{k-1} \{F_n > -\theta\} \middle| t_1 > k\right] P[t_1 > k]
\end{aligned}
$$

Comparison of Analysis with Simulations

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)

Comparison of Analysis with Simulations

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0, 1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

Comparison of Analysis with Simulations

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

$P_{MD} Sim.$	$P_{MD} Anal.$	$E_{DD} Sim.$	$E_{DD} Anal.$
$18.78e-4$	$19.85e-4$	44.319	43.290
$26.68e-4$	$27.51e-4$	36.028	34.634
$36.30e-4$	$35.16e-4$	27.770	25.977

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{reject H_1\}]$$

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{reject H_1\}]$$

$$D_{tot}^0 = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D_{tot}^1 = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D_{tot}^0},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D_{tot}^1}$$

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{reject H_1\}]$$

$$D^0_{tot} = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D^1_{tot} = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D^0_{tot}},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D^1_{tot}}$$

Theorem

- $\gamma_{0,l} = -r_l|\log c|\ , \gamma_{1,l} = \rho_l|\log c|,\ \beta_0 = -|\log c|,\ \beta_1 = |\log c|.$

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{rejectH_1\}]$$

$$D_{tot}^0 = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D_{tot}^1 = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D_{tot}^0},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D_{tot}^1}$$

Theorem

- $\gamma_{0,l} = -r_l|\log c|\ ,\gamma_{1,l} = \rho_l|\log c|,\ \beta_0 = -|\log c|,\ \beta_1 = |\log c|.$

$$\lim_{c \to 0} \frac{E_0[N]}{|\log c|} \leq \frac{1}{D_{tot}^0} + M_0 \text{ and } \lim_{c \to 0} \frac{E_1[N]}{|\log c|} \leq \frac{1}{D_{tot}^1} + M_1$$

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{rejectH_1\}]$$

$$D^0_{tot} = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D^1_{tot} = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D^0_{tot}},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D^1_{tot}}$$

Theorem

- $\gamma_{0,l} = -r_l|\log c|\ ,\gamma_{1,l} = \rho_l|\log c|,\ \beta_0 = -|\log c|,\ \beta_1 = |\log c|.$

$$\lim_{c\to 0} \frac{E_0[N]}{|\log c|} \le \frac{1}{D^0_{tot}} + M_0 \text{ and } \lim_{c\to 0} \frac{E_1[N]}{|\log c|} \le \frac{1}{D^1_{tot}} + M_1$$

- Gaussian MAC noise at FC
- μ chosen appropriately

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{reject H_1\}]$$

$$D^0_{tot} = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D^1_{tot} = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D^0_{tot}},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D^1_{tot}}$$

Theorem

- $\gamma_{0,l} = -r_l|\log c|\ ,\gamma_{1,l} = \rho_l|\log c|,\ \beta_0 = -|\log c|,\ \beta_1 = |\log c|.$

$$\lim_{c\to 0} \frac{E_0[N]}{|\log c|} \le \frac{1}{D^0_{tot}} + M_0 \text{ and } \lim_{c\to 0} \frac{E_1[N]}{|\log c|} \le \frac{1}{D^1_{tot}} + M_1$$

- Gaussian MAC noise at FC
- μ chosen appropriately

$$\lim_{c\to 0} \frac{P_{FA}}{c|\log c|} = 0 \text{ and } \lim_{c\to 0} \frac{P_{MD}}{c|\log c|} = 0$$

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\,E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\,E_1(N) + W_1 P_1\{rejectH_1\}]$$

$$D_{tot}^0 = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D_{tot}^1 = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D_{tot}^0},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D_{tot}^1}$$

Theorem

- $\gamma_{0,l} = -r_l|\log c|\ ,\gamma_{1,l} = \rho_l|\log c|,\ \beta_0 = -|\log c|,\ \beta_1 = |\log c|.$

$$\lim_{c\to 0} \frac{E_0[N]}{|\log c|} \le \frac{1}{D_{tot}^0} + M_0 \text{ and } \lim_{c\to 0} \frac{E_1[N]}{|\log c|} \le \frac{1}{D_{tot}^1} + M_1$$

- Gaussian MAC noise at FC
- μ chosen appropriately

$$\lim_{c\to 0} \frac{P_{FA}}{c|\log c|} = 0 \text{ and } \lim_{c\to 0} \frac{P_{MD}}{c|\log c|} = 0$$

$$\lim_{c\to 0} \frac{\mathcal{R}_c(\delta^*_{Bayes})}{\mathcal{R}_c(\delta)} = 1$$

Asymptotic Properties of DualSPRT

$$\mathcal{R}_c(\delta) = \pi[c\, E_0(N) + W_0 P_0\{reject\ \ H_0\}] + (1-\pi)[c\, E_1(N) + W_1 P_1\{rejectH_1\}]$$

$$D_{tot}^0 = \sum_{l=1}^{L} D(f_{0,l}||f_{1,l}),\ D_{tot}^1 = \sum_{l=1}^{L} D(f_{1,l}||f_{0,l}),\ r_l = \frac{D(f_{0,l}||f_{1,l})}{D_{tot}^0},\ \rho_l = \frac{D(f_{1,l}||f_{0,l})}{D_{tot}^1}$$

Theorem

- $\gamma_{0,l} = -r_l|\log c|\ ,\gamma_{1,l} = \rho_l|\log c|,\ \beta_0 = -|\log c|,\ \beta_1 = |\log c|.$

$$\lim_{c\to 0} \frac{E_0[N]}{|\log c|} \leq \frac{1}{D_{tot}^0} + M_0 \text{ and } \lim_{c\to 0} \frac{E_1[N]}{|\log c|} \leq \frac{1}{D_{tot}^1} + M_1$$

- Gaussian MAC noise at FC
- μ chosen appropriately

$$\lim_{c\to 0} \frac{P_{FA}}{c|\log c|} = 0 \text{ and } \lim_{c\to 0} \frac{P_{MD}}{c|\log c|} = 0$$

$$\lim_{c\to 0} \frac{\mathcal{R}_c(\delta^*_{Bayes})}{\mathcal{R}_c(\delta)} = 1 \text{ and } \lim_{c\to 0} \frac{\mathcal{R}_c(\delta^*_{Cent.})}{\mathcal{R}_c(\delta)} = 1$$

Modification to DualSPRT: SPRT-CSPRT Algorithm

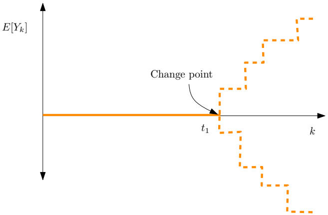
- Using SPRT at Fusion center is not optimal since the optimality of SPRT is known for i.i.d. observations only.

Modification to DualSPRT: SPRT-CSPRT Algorithm

- Using SPRT at Fusion center is not optimal since the optimality of SPRT is known for i.i.d. observations only.
- What is the best test at FC?

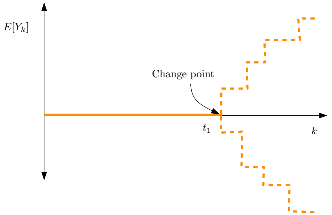
Modification to DualSPRT: SPRT-CSPRT Algorithm

- Using SPRT at Fusion center is not optimal since the optimality of SPRT is known for i.i.d. observations only.
- What is the best test at FC?
- Sequential change detection formulation $\mu_k = E[Y_k]$



Modification to DualSPRT: SPRT-CSPRT Algorithm

- Using SPRT at Fusion center is not optimal since the optimality of SPRT is known for i.i.d. observations only.
- What is the best test at FC?
- Sequential change detection formulation $\mu_k = E[Y_k]$



- Reducing the false alarms caused by FC MAC noise before t_1.

- SPRT-CSPRT: (inspired from CUSUM) FC algorithm,

$$F_k^1 = \left(F_{k-1}^1 + \log \frac{g_{\mu_1}(Y_k)}{g_z(Y_k)}\right)^+, \; F_k^0 = \left(F_{k-1}^0 + \log \frac{g_z(Y_k)}{g_{-\mu_0}(Y_k)}\right)^-$$

- SPRT-CSPRT: (inspired from CUSUM) FC algorithm,

$$F_k^1 = \left(F_{k-1}^1 + \log \frac{g_{\mu_1}(Y_k)}{g_z(Y_k)}\right)^+, \; F_k^0 = \left(F_{k-1}^0 + \log \frac{g_z(Y_k)}{g_{-\mu_0}(Y_k)}\right)^-$$

- FC decides at $N = \inf\{k : F_k^1 \geq \beta_1 \; or \; F_k^0 \leq -\beta_0\}$,

$$H_1 \text{ if } F_N^1 \geq \beta_1, H_0 \text{ if } F_N^0 \leq -\beta_0$$

- SPRT-CSPRT: (inspired from CUSUM) FC algorithm,

$$F_k^1 = \left(F_{k-1}^1 + \log \frac{g_{\mu_1}(Y_k)}{g_z(Y_k)}\right)^+, \; F_k^0 = \left(F_{k-1}^0 + \log \frac{g_z(Y_k)}{g_{-\mu_0}(Y_k)}\right)^-$$

- FC decides at $N = \inf\{k : F_k^1 \geq \beta_1 \; or \; F_k^0 \leq -\beta_0\}$,

$$H_1 \text{ if } F_N^1 \geq \beta_1, H_0 \text{ if } F_N^0 \leq -\beta_0$$

- $\widehat{\mu}_k = E[Y_k], \; E_{\widehat{\mu}_k}\left[\log \frac{g_{\mu_1}(Y_k)}{g_Z(Y_k)}\right] = D(g_{\widehat{\mu}_k}||g_Z) - D(g_{\widehat{\mu}_k}||g_{\mu_1})$

- SPRT-CSPRT: (inspired from CUSUM) FC algorithm,

$$F_k^1 = \left(F_{k-1}^1 + \log \frac{g_{\mu_1}(Y_k)}{g_z(Y_k)} \right)^+, \; F_k^0 = \left(F_{k-1}^0 + \log \frac{g_z(Y_k)}{g_{-\mu_0}(Y_k)} \right)^-$$

- FC decides at $N = \inf\{k : F_k^1 \geq \beta_1 \; or \; F_k^0 \leq -\beta_0\}$,

$$H_1 \text{ if } F_N^1 \geq \beta_1, H_0 \text{ if } F_N^0 \leq -\beta_0$$

- $\widehat{\mu}_k = E[Y_k]$, $E_{\widehat{\mu}_k}\left[\log \frac{g_{\mu_1}(Y_k)}{g_Z(Y_k)}\right] = D(g_{\widehat{\mu}_k} || g_Z) - D(g_{\widehat{\mu}_k} || g_{\mu_1})$
  Exp. drift of F_k^1 +ve if $D(g_{\widehat{\mu}_k} || g_Z) > D(g_{\widehat{\mu}_k} || g_{\mu_1})$.

- SPRT-CSPRT: (inspired from CUSUM) FC algorithm,

$$F_k^1 = \left(F_{k-1}^1 + \log \frac{g_{\mu_1}(Y_k)}{g_z(Y_k)}\right)^+, \; F_k^0 = \left(F_{k-1}^0 + \log \frac{g_z(Y_k)}{g_{-\mu_0}(Y_k)}\right)^-$$

- FC decides at $N = \inf\{k : F_k^1 \geq \beta_1 \text{ } or \text{ } F_k^0 \leq -\beta_0\}$,

$$H_1 \text{ if } F_N^1 \geq \beta_1, H_0 \text{ if } F_N^0 \leq -\beta_0$$

- $\widehat{\mu}_k = E[Y_k]$, $E_{\widehat{\mu}_k}\left[\log \frac{g_{\mu_1}(Y_k)}{g_Z(Y_k)}\right] = D(g_{\widehat{\mu}_k}||g_Z) - D(g_{\widehat{\mu}_k}||g_{\mu_1})$
Exp. drift of F_k^1 $+$ve if $D(g_{\widehat{\mu}_k}||g_Z) > D(g_{\widehat{\mu}_k}||g_{\mu_1})$.
Exp. drift of F_k^0 -ve if $D(g_{\widehat{\mu}_k}||g_{-\mu_0}) < D(g_{\widehat{\mu}_k}||g_Z)$.

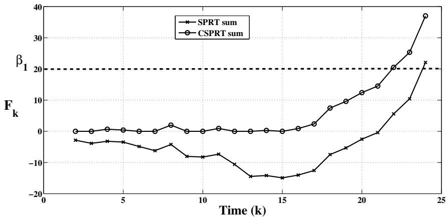
- SPRT-CSPRT: (inspired from CUSUM) FC algorithm,

$$F_k^1 = \left(F_{k-1}^1 + \log \frac{g_{\mu_1}(Y_k)}{g_z(Y_k)}\right)^+, \; F_k^0 = \left(F_{k-1}^0 + \log \frac{g_z(Y_k)}{g_{-\mu_0}(Y_k)}\right)^-$$

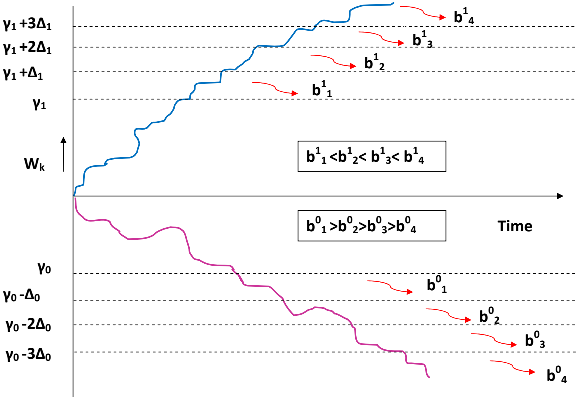
- FC decides at $N = \inf\{k : F_k^1 \geq \beta_1 \; or \; F_k^0 \leq -\beta_0\}$,

$$H_1 \text{ if } F_N^1 \geq \beta_1, H_0 \text{ if } F_N^0 \leq -\beta_0$$

- $\widehat{\mu}_k = E[Y_k]$, $E_{\widehat{\mu}_k}\left[\log \frac{g_{\mu_1}(Y_k)}{g_Z(Y_k)}\right] = D(g_{\widehat{\mu}_k}||g_Z) - D(g_{\widehat{\mu}_k}||g_{\mu_1})$
  Exp. drift of F_k^1 +ve if $D(g_{\widehat{\mu}_k}||g_Z) > D(g_{\widehat{\mu}_k}||g_{\mu_1})$.
  Exp. drift of F_k^0 -ve if $D(g_{\widehat{\mu}_k}||g_{-\mu_0}) < D(g_{\widehat{\mu}_k}||g_Z)$.
- Sample path argument



Modification of quantisation at SU's



Performance Analysis of SPRT-CSPRT: P_{MD}

- $P_{H_1}(MD) = P_{H_1}(MD < t_1) + P_{H_1}(t_1 \leq MD \leq t_2) + \ldots$
- First term dominates.

Performance Analysis of SPRT-CSPRT: P_{MD}

- $P_{H_1}(MD) = P_{H_1}(MD < t_1) + P_{H_1}(t_1 \leq MD \leq t_2) + \ldots$
- First term dominates.

$$
\begin{aligned}
P_{H_1}(\text{FA before } t_1) &= \sum_{k=1}^{\infty} P(\tau_\beta \leq k | k < t_1) P(t_1 > k) \\
\tau_\beta &\triangleq \inf\{k \geq 1 : F_k^0 \leq -\beta\}
\end{aligned}
$$

Performance Analysis of SPRT-CSPRT: P_{MD}

- $P_{H_1}(MD) = P_{H_1}(MD < t_1) + P_{H_1}(t_1 \leq MD \leq t_2) + \dots$
- First term dominates.

$$
\begin{aligned}
P_{H_1}(\text{FA before } t_1) &= \sum_{k=1}^{\infty} P(\tau_\beta \leq k | k < t_1) P(t_1 > k) \\
\tau_\beta &\triangleq \inf\{k \geq 1 : F_k^0 \leq -\beta\} \\
\lim_{\beta \to \infty} P_0\{\tau_\beta > x | \tau_\beta < t_1\} &= \exp(-\lambda_\beta x), x > 0
\end{aligned}
$$

Performance Analysis of SPRT-CSPRT: P_{MD}

- $P_{H_1}(MD) = P_{H_1}(MD < t_1) + P_{H_1}(t_1 \le MD \le t_2) + \dots$
- First term dominates.

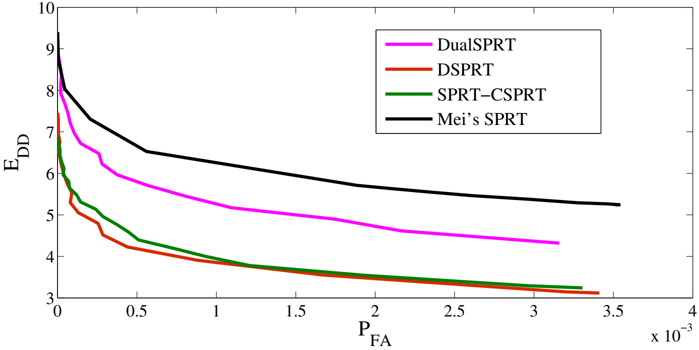
$$
\begin{aligned}
P_{H_1}(\text{FA before } t_1) &= \sum_{k=1}^{\infty} P(\tau_\beta \le k | k < t_1) P(t_1 > k) \\
\tau_\beta &\triangleq \inf\{k \ge 1 : F_k^0 \le -\beta\} \\
\lim_{\beta\to\infty} P_0\{\tau_\beta > x | \tau_\beta < t_1\} &= \exp(-\lambda_\beta x), x > 0 \\
P_{H_1}(\text{FA before } t_1) &\approx \sum_{k=1}^{\infty} (1 - e^{-\lambda_\beta k}) \prod_{l=1}^{L} (1 - \Phi_{\tau_{\gamma,l}}(k))
\end{aligned}
$$

Comparison of Analysis with Simulations

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

$P_{MD} Sim.$	$P_{MD} Anal.$	$E_{DD} Sim.$	$E_{DD} Anal.$
0.00675	0.00613	26.8036	24.9853
0.0072	0.0065	33.1585	31.7624
0.01675	0.01624	30.0817	29.1322

Comparison with asymptotically optimal tests



Mei's SPRT: Mei, Trans. Inf. Theory 2008.

DSPRT: Fellouris and Moustakides, Trans. Inf. Theory 2011.

Different and unknown SNR's: GLR-SPRT

- SNR at local CR nodes Unknown

Different and unknown SNR's: GLR-SPRT

- SNR at local CR nodes Unknown
- Energy detector output is used as the observations to the local node SPRT's

Different and unknown SNR's: GLR-SPRT

- SNR at local CR nodes Unknown
- Energy detector output is used as the observations to the local node SPRT's
- Modelled as Gaussian mean change under low SNR's and large number of samples

Different and unknown SNR's: GLR-SPRT

- SNR at local CR nodes Unknown
- Energy detector output is used as the observations to the local node SPRT's
- Modelled as Gaussian mean change under low SNR's and large number of samples
- At SU l,

$$H_0 : \theta = \theta_0 \ ; \ H_1 : \theta \geq \theta_1 \ .$$

where $\theta_0 = 0$ and θ_1 is appropriately chosen,

Different and unknown SNR's: GLR-SPRT

- SNR at local CR nodes Unknown
- Energy detector output is used as the observations to the local node SPRT's
- Modelled as Gaussian mean change under low SNR's and large number of samples
- At SU l,

$$H_0 : \theta = \theta_0 \ ; \ H_1 : \theta \geq \theta_1 \ .$$

where $\theta_0 = 0$ and θ_1 is appropriately chosen,

$$W_{n,l} = \max\left[\sum_{k=1}^{n} \log\frac{f_{\hat{\theta}_n}(X_k)}{f_{\theta_0}(X_k)}, \sum_{k=1}^{n} \log\frac{f_{\hat{\theta}_n}(X_k)}{f_{\theta_1}(X_k)}\right] \ ,$$

Different and unknown SNR's: GLR-SPRT

- SNR at local CR nodes Unknown
- Energy detector output is used as the observations to the local node SPRT's
- Modelled as Gaussian mean change under low SNR's and large number of samples
- At SU l,

$$H_0 : \theta = \theta_0 \ ; \ H_1 : \theta \geq \theta_1 \ .$$

where $\theta_0 = 0$ and θ_1 is appropriately chosen,

$$W_{n,l} = \max \left[\sum_{k=1}^{n} \log \frac{f_{\hat{\theta}_n}(X_k)}{f_{\theta_0}(X_k)}, \sum_{k=1}^{n} \log \frac{f_{\hat{\theta}_n}(X_k)}{f_{\theta_1}(X_k)} \right] \ ,$$

$$\begin{aligned} N &= \inf \{ n : W_{n,l} \geq g(cn) \} \, , \\ g(t) &\approx \log(1/t) \text{ as } t \to 0 \end{aligned}$$

- Decide upon H_0 or H_1 as $\hat{\theta}_N \leq \theta^*$ *or* $\hat{\theta}_N \geq \theta^*$ where θ^* solves $D(f_{\theta^*}||f_{\theta_0}) = D(f_{\theta^*}||f_{\theta_1})$
- $D(f_\theta||f_\lambda) = E_\theta[\log[f_\theta(X)/f_\lambda(X)]$

- Decide upon H_0 or H_1 as $\hat{\theta}_N \leq \theta^* \; or \; \hat{\theta}_N \geq \theta^*$ where θ^* solves $D(f_{\theta^*}||f_{\theta_0}) = D(f_{\theta^*}||f_{\theta_1})$
- $D(f_\theta||f_\lambda) = E_\theta[\log[f_\theta(X)/f_\lambda(X)]$
- Similar to GLR test in Neyman-Pearson approach in FSS tests
- Diminishing uncertainty of $\hat{\theta}_n$ as an estimate of θ with n is incorporated into the varying stopping boundary $g(cn)$
- Asymptotically optimal over a broad range of θ as $c \to 0$

Comparison between DualSPRT and GLRSPRT

- 5 Secondary nodes

Comparison between DualSPRT and GLRSPRT

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)

Comparison between DualSPRT and GLRSPRT

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

Comparison between DualSPRT and GLRSPRT

- 5 Secondary nodes
- channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0, 1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

hyp	E_{DD}	$P_E = 0.1$	$P_E = 0.05$	$P_E = 0.01$
H1	DualSPRT	2.06	3.177	5.264
H1	GLRSPRT	1.425	2.522	4.857
H0	DualSPRT	1.921	3.074	5.184
H0	GLRSPRT	2.745	3.852	6.115

Rayleigh Slow fading case

- N is much less than the coherence time of the channel

Rayleigh Slow fading case

- N is much less than the coherence time of the channel
- Fading gain is unknown to the CR nodes

Rayleigh Slow fading case

- N is much less than the coherence time of the channel
- Fading gain is unknown to the CR nodes
- In case of Rayleigh fading, the received signal power, P_I is exponentially distributed

Rayleigh Slow fading case

- N is much less than the coherence time of the channel
- Fading gain is unknown to the CR nodes
- In case of Rayleigh fading, the received signal power, P_l is exponentially distributed

$$H_0 : f_{0,l} \sim \mathcal{N}(0, \sigma^2); H_1 : f_{1,l} \sim \mathcal{N}(\theta, \sigma^2)$$

where θ is exponential distributed r.v which reflects the unknown channel gain.

Simulation results

- No. of nodes :5
- $\sigma^2 = 1, \theta = exp(1), Var(Z_k) = 1$

Simulation results

- No. of nodes :5
- $\sigma^2 = 1, \theta = exp(1), Var(Z_k) = 1$

E_{DD}	$P_{FA} = 0.1$	$P_{FA} = 0.05$	$P_{FA} = 0.01$
DualSPRT	1.669	2.497	4.753
GLRSPRT	3.191	4.418	7.294

Table: slow-fading between primary and secondary user under H0

Simulation results

- No. of nodes :5
- $\sigma^2 = 1, \theta = exp(1), Var(Z_k) = 1$

E_{DD}	$P_{FA} = 0.1$	$P_{FA} = 0.05$	$P_{FA} = 0.01$
DualSPRT	1.669	2.497	4.753
GLRSPRT	3.191	4.418	7.294

Table: slow-fading between primary and secondary user under H0

E_{DD}	$P_{MD} = 0.1$	$P_{MD} = 0.08$	$P_{MD} = 0.06$
DualSPRT	1.74	1.854	2.417
GLRSPRT	1.62	3.065	5.42

Table: slow-fading between primary and secondary user under H1

Comparison between SPRT-CSPRT and GLR-CSPRT

- 5 Secondary nodes
- Unknown SNR: channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0, 1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

Comparison between SPRT-CSPRT and GLR-CSPRT

- 5 Secondary nodes
- Unknown SNR: channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

Hyp	E_{DD}	$P_E = 0.1$	$P_E = 0.05$	$P_E = 0.01$
H_1	SPRT-CSPRT	1.615	2.480	4.28
H_1	GLR-CSPRT	1.138	2.221	4.533
H_0	SPRT-CSPRT	1.533	2.334	4.225
H_0	GLR-CSPRT	2.424	3.734	5.72

Comparison between SPRT-CSPRT and GLR-CSPRT

- 5 Secondary nodes
- Unknown SNR: channel gains (0, -1.5, -2.5, -4 and -6 dB)
- $f_0 \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(\theta_l, 1)$, $Var(Z_k) = 1$

Hyp	E_{DD}	$P_E = 0.1$	$P_E = 0.05$	$P_E = 0.01$
H_1	SPRT-CSPRT	1.615	2.480	4.28
H_1	GLR-CSPRT	1.138	2.221	4.533
H_0	SPRT-CSPRT	1.533	2.334	4.225
H_0	GLR-CSPRT	2.424	3.734	5.72

- Rayleigh Slow fading: $\sigma^2 = 1, \theta = exp(1), Var(Z_k) = 1$

Hyp	E_{DD}	$P_E = 0.1$	$P_E = 0.07$	$P_E = 0.04$
H_1	SPRT-CSPRT	1.03	1.53	2.347
H_1	GLR-CSPRT	0.94	1.004	4.225
H_0	SPRT-CSPRT	1.528	1.741	2.415
H_0	GLR-CSPRT	2.615	3.192	4.237

Universal Sequential Hypothesis Testing

- Nonparametric or universal setup:
  - Noise statistics under no PU transmission is fully known
  - Transmit power, channel gains, modulation schemes etc. of PU transmissions is not available (SNR uncertainty).
- i.i.d. observations $X_i, i = 1, 2, \ldots$
- $H_0 : X_i \sim P_0($ p.d.f. $f_0)$ known ;
  $H_1 : X_i \sim P_1($ p.d.f. $f_1)$ unknown

Discrete alphabet for P_0 and P_1

- SPRT: P_0 and P_1 fully known, for $\gamma_0, \gamma_1 > 0$

Discrete alphabet for P_0 and P_1

- SPRT: P_0 and P_1 fully known, for $\gamma_0, \gamma_1 > 0$

$$N \triangleq \inf\{n : W_n = \sum_{k=1}^{n} \log \frac{P_1(X_k)}{P_0(X_k)} \notin (-\gamma_0, \gamma_1)\},$$

Discrete alphabet for P_0 and P_1

- SPRT: P_0 and P_1 fully known, for $\gamma_0, \gamma_1 > 0$

$$
\begin{aligned}
N &\triangleq \inf\left\{n : W_n = \sum_{k=1}^{n} \log \frac{P_1(X_k)}{P_0(X_k)} \notin (-\gamma_0, \gamma_1)\right\}, \\
\delta &= H_1 \text{ if } W_N \geq \gamma_1;\ H_0 \text{ if } W_N \leq -\gamma_0
\end{aligned}
$$

Discrete alphabet for P_0 and P_1

- SPRT: P_0 and P_1 fully known, for $\gamma_0, \gamma_1 > 0$

$$
\begin{aligned}
N &\triangleq \inf\{n : W_n = \sum_{k=1}^{n} \log \frac{P_1(X_k)}{P_0(X_k)} \notin (-\gamma_0, \gamma_1)\}, \\
\delta &= H_1 \text{ if } W_N \geq \gamma_1;\ H_0 \text{ if } W_N \leq -\gamma_0
\end{aligned}
$$

- P_0 is known; P_1 is not known.

Discrete alphabet for P_0 and P_1

- SPRT: P_0 and P_1 fully known, for $\gamma_0, \gamma_1 > 0$

$$\begin{aligned} N &\triangleq \inf\{n : W_n = \sum_{k=1}^{n} \log \frac{P_1(X_k)}{P_0(X_k)} \notin (-\gamma_0, \gamma_1)\}, \\ \delta &= H_1 \text{ if } W_N \geq \gamma_1;\ H_0 \text{ if } W_N \leq -\gamma_0 \end{aligned}$$

- P_0 is known; P_1 is not known. Replace W_n by

$$\widehat{W}_n = -L_n(X_1^n) - \log P_0(X_1^n) - n\frac{\lambda}{2},\ \lambda > 0.$$

$L_n(X_1^n)$=Codelength of a universal lossless source code for X_1^n

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1} \log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1}\log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).
- Consider universal codes which satisfy $\lim_{n\to\infty} n^{-1} L_n = \bar{H}(X)$ a.s. atleast for i.i.d. sources.

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1} \log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).
- Consider universal codes which satisfy $\lim_{n\to\infty} n^{-1} L_n = \bar{H}(X)$ a.s. atleast for i.i.d. sources.
- Under H_1:
  - $E_1[(-\log P_0(X_1^n))] = nH_1(X) + nD(P_1||P_0)$

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1} \log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).
- Consider universal codes which satisfy $\lim_{n\to\infty} n^{-1} L_n = \bar{H}(X)$ a.s. atleast for i.i.d. sources.
- Under H_1:
  - $E_1[(-\log P_0(X_1^n))] = nH_1(X) + nD(P_1||P_0)$
  - $L(X_1^n) \approx nH_1(X)$ for large n

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1}\log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).
- Consider universal codes which satisfy $\lim_{n\to\infty} n^{-1}L_n = \bar{H}(X)$ a.s. atleast for i.i.d. sources.
- Under H_1:
  - $E_1[(-\log P_0(X_1^n))] = nH_1(X) + nD(P_1||P_0)$
  - $L(X_1^n) \approx nH_1(X)$ for large n
  - Thus average drift under H_1: $D(P_1||P_0) - \lambda/2$

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1}\log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).
- Consider universal codes which satisfy $\lim_{n\to\infty} n^{-1} L_n = \bar{H}(X)$ a.s. atleast for i.i.d. sources.
- Under H_1:
  - $E_1[(-\log P_0(X_1^n))] = nH_1(X) + nD(P_1||P_0)$
  - $L(X_1^n) \approx nH_1(X)$ for large n
  - Thus average drift under H_1: $D(P_1||P_0) - \lambda/2$

  Average drift under H_0: $-\lambda/2$

- By Shanon-Macmillan Thorem, for any stationary ergodic source $\lim_{n\to\infty} n^{-1} \log P(X_1^n) = -\bar{H}(X)$ a.s. ($\bar{H}(X)$ is the entropy rate).
- Consider universal codes which satisfy $\lim_{n\to\infty} n^{-1} L_n = \bar{H}(X)$ a.s. atleast for i.i.d. sources.
- Under H_1:
  - $E_1[(-\log P_0(X_1^n))] = nH_1(X) + nD(P_1||P_0)$
  - $L(X_1^n) \approx nH_1(X)$ for large n
  - Thus average drift under H_1: $D(P_1||P_0) - \lambda/2$

  Average drift under H_0: $-\lambda/2$

  Thus consider P_1 belonging to the class

$$\mathcal{C} = \{P_1 : D(P_1, P_0) \geq \lambda\}.$$

(A): stronger version of pointwise universality

Theorem

(1) $P_0(N < \infty) = 1$ and $P_1(N < \infty) = 1$.

(A): stronger version of pointwise universality

Theorem

(1) $P_0(N < \infty) = 1$ and $P_1(N < \infty) = 1$.

(2) $P_{FA} \triangleq P_0(\widehat{W}_N \geq \gamma_1) \leq \exp(-\gamma_1)$.

Discrete alphabet case
Asymptotic Properties

(A): stronger version of pointwise universality

Theorem

(1) $P_0(N < \infty) = 1$ and $P_1(N < \infty) = 1$.

(2) $P_{FA} \triangleq P_0(\widehat{W}_N \geq \gamma_1) \leq \exp(-\gamma_1)$.

(3) Under (A), $P_{MD} \triangleq P_1(\widehat{W}_N \leq -\gamma_0) = \mathcal{O}(\exp(-\gamma_0 s)),\ s > 0.$.

(A): stronger version of pointwise universality

Theorem

(1) $P_0(N < \infty) = 1$ and $P_1(N < \infty) = 1$.

(2) $P_{FA} \triangleq P_0(\widehat{W}_N \geq \gamma_1) \leq \exp(-\gamma_1)$.

(3) Under (A), $P_{MD} \triangleq P_1(\widehat{W}_N \leq -\gamma_0) = \mathcal{O}(\exp(-\gamma_0 s)),\ s > 0.$.

(4) $\dfrac{N}{|\log \gamma_0|} \xrightarrow[\gamma_0,\gamma_1 \to \infty]{P_0 \text{ a.s.}} \dfrac{2}{\lambda};\quad \dfrac{N}{|\log \gamma_1|} \xrightarrow[\gamma_0,\gamma_1 \to \infty]{P_1,\text{ a.s.}} \dfrac{1}{D(P_1||P_0) - \lambda/2}$

(A): stronger version of pointwise universality

Theorem

(1) $P_0(N < \infty) = 1$ and $P_1(N < \infty) = 1$.

(2) $P_{FA} \triangleq P_0(\widehat{W}_N \geq \gamma_1) \leq \exp(-\gamma_1)$.

(3) Under (A), $P_{MD} \triangleq P_1(\widehat{W}_N \leq -\gamma_0) = \mathcal{O}(\exp(-\gamma_0 s)),\ s > 0.$.

(4) $\dfrac{N}{|\log \gamma_0|} \xrightarrow[\gamma_0,\gamma_1 \to \infty]{P_0 \text{ a.s.}} \dfrac{2}{\lambda}$; $\dfrac{N}{|\log \gamma_1|} \xrightarrow[\gamma_0,\gamma_1 \to \infty]{P_1, \text{ a.s.}} \dfrac{1}{D(P_1||P_0) - \lambda/2}$

(5) Under (A), if $E_1[(\log P_1(X_1))^{p+1}] < \infty$ & $E_1[(\log P_0(X_1))^{p+1}] < \infty$ for some $p \geq 1$, then for all $0 < q \leq p$,

$$\frac{E_0[N^q]}{|\log \gamma_0|^q} \xrightarrow[\gamma_0,\gamma_1 \to \infty]{} \left(\frac{2}{\lambda}\right)^q, \quad \frac{E_1[N^q]}{|\log \gamma_1|^q} \xrightarrow[\gamma_0,\gamma_1 \to \infty]{} \left(\frac{1}{D(P_1||P_0) - \frac{\lambda}{2}}\right)^q$$

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$
  If $E[\log f_1(X_k)] < \infty$ then by SLLN

$$n^{-1} \sum_{k=1}^{n} \log f_1(X_k) \to -h(X_1) \text{ a.s.}$$

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$
  If $E[\log f_1(X_k)] < \infty$ then by SLLN

$$n^{-1} \sum_{k=1}^{n} \log f_1(X_k) \to -h(X_1) \text{ a.s.}$$

- $X_k \to X_k^{\Delta} = [X_k/\Delta]\Delta;$

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$
  If $E[\log f_1(X_k)] < \infty$ then by SLLN

$$n^{-1} \sum_{k=1}^{n} \log f_1(X_k) \to -h(X_1) \text{ a.s.}$$

- $X_k \to X_k^{\Delta} = [X_k/\Delta]\Delta$;
  Use Universal lossless coding on $X_1^{\Delta}, X_2^{\Delta}, \ldots, X_n^{\Delta}$

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$
  If $E[\log f_1(X_k)] < \infty$ then by SLLN

$$n^{-1} \sum_{k=1}^{n} \log f_1(X_k) \to -h(X_1) \text{ a.s.}$$

- $X_k \to X_k^{\Delta} = [X_k/\Delta]\Delta$;
  Use Universal lossless coding on $X_1^{\Delta}, X_2^{\Delta}, \ldots, X_n^{\Delta}$
- $H(X_1^{\Delta}) + \log \Delta \to h(X_1)$ as $\Delta \to 0$.

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$
  If $E[\log f_1(X_k)] < \infty$ then by SLLN

$$n^{-1} \sum_{k=1}^{n} \log f_1(X_k) \to -h(X_1) \text{ a.s.}$$

- $X_k \to X_k^\Delta = [X_k/\Delta]\Delta$;
  Use Universal lossless coding on $X_1^\Delta, X_2^\Delta, \ldots, X_n^\Delta$
- $H(X_1^\Delta) + \log \Delta \to h(X_1)$ as $\Delta \to 0$.
- Test is modified to

$$\widetilde{W}_n = -L_n(X_{1:n}^\Delta) - n \log \Delta - \sum_{k=1}^{n} \log f_0(X_k) - n\frac{\lambda}{2}$$

Continuous alphabet for P_0 and P_1

- Need to estimate $\sum_{k=1}^{n} \log f_1(X_k)$
  If $E[\log f_1(X_k)] < \infty$ then by SLLN

$$n^{-1} \sum_{k=1}^{n} \log f_1(X_k) \to -h(X_1) \text{ a.s.}$$

- $X_k \to X_k^{\Delta} = [X_k/\Delta]\Delta$;
  Use Universal lossless coding on $X_1^{\Delta}, X_2^{\Delta}, \ldots, X_n^{\Delta}$
- $H(X_1^{\Delta}) + \log \Delta \to h(X_1)$ as $\Delta \to 0$.
- Test is modified to

$$\widetilde{W}_n = -L_n(X_{1:n}^{\Delta}) - n \log \Delta - \sum_{k=1}^{n} \log f_0(X_k) - n\frac{\lambda}{2}$$

$$f_1 \in \mathcal{C} = \{f_1 : D(f_1, f_0) \geq \lambda\}.$$

- Uniform scalar quantization with optimal source coding is optimal at high rates.

- Uniform scalar quantization with optimal source coding is optimal at high rates.
- An adaptive uniform quantizer makes $\{X_n^{\Delta}\}$ non i.i.d. L_n unable to learn the underlying distribution in such scenario.

- Uniform scalar quantization with optimal source coding is optimal at high rates.
- An adaptive uniform quantizer makes $\{X_n^{\Delta}\}$ non i.i.d. L_n unable to learn the underlying distribution in such scenario.
- Non-uniform partitions with Δ_j in j^{th} bin and with p.m.f p_j require knowledge of p_j which is unknown under H_1.

- Uniform scalar quantization with optimal source coding is optimal at high rates.
- An adaptive uniform quantizer makes $\{X_n^{\Delta}\}$ non i.i.d. L_n unable to learn the underlying distribution in such scenario.
- Non-uniform partitions with Δ_j in j^{th} bin and with p.m.f p_j require knowledge of p_j which is unknown under H_1.
- $f_0(X_i)\Delta \approx p_0(X_i^{\Delta})$. $\widetilde{W}_n = -L_n(X_i^{\Delta}) - \sum_{j=1}^{n} \log p_0(X_i^{\Delta}) - n\lambda/2$.

Continuous alphabet case

Why Uniform Quantization?

- Uniform scalar quantization with optimal source coding is optimal at high rates.
- An adaptive uniform quantizer makes $\{X_n^{\Delta}\}$ non i.i.d. L_n unable to learn the underlying distribution in such scenario.
- Non-uniform partitions with Δ_j in j^{th} bin and with p.m.f p_j require knowledge of p_j which is unknown under H_1.
- $f_0(X_i)\Delta \approx p_0(X_i^{\Delta})$. $\widetilde{W}_n = -L_n(X_i^{\Delta}) - \sum_{j=1}^{n} \log p_0(X_i^{\Delta}) - n\lambda/2$.
- Range of the quantization: f_1's tail probabilities less than a small specific value at a fixed boundary

LZSLRT algorithm

- Use Lempel-Ziv incremental parsing technique (LZ78).

LZSLRT algorithm

- Use Lempel-Ziv incremental parsing technique (LZ78).
- Added a correction term $n\epsilon_n$ in $\widetilde{W}_n$

LZSLRT algorithm

- Use Lempel-Ziv incremental parsing technique (LZ78).
- Added a correction term $n\epsilon_n$ in $\widetilde{W}_n$

$$\epsilon_n = C\left(\frac{1}{\log n} + \frac{\log\log n}{n} + \frac{\log\log n}{\log n}\right).$$

C depends on size of the quantized alphabet.

LZSLRT algorithm

- Use Lempel-Ziv incremental parsing technique (LZ78).
- Added a correction term $n\epsilon_n$ in $\widetilde{W}_n$

$$\epsilon_n = C\left(\frac{1}{\log n} + \frac{\log\log n}{n} + \frac{\log\log n}{\log n}\right).$$

  C depends on size of the quantized alphabet.
- LZSLRT performance is close to the nearly optimal parametric GLR sequential tests and is better for some class of distributions (e.g. Pareto)

KTSLRT algorithm

- $L_n(X_{1:n}^{\Delta})$ from combined use of Krichevsky-Trofimov estimator and the Arithmetic Encoder

KTSLRT algorithm

- $L_n(X^{\Delta}_{1:n})$ from combined use of Krichevsky-Trofimov estimator and the Arithmetic Encoder
  - Estimate source distribution via K-T estimator for a finite alphabet source as,

$$P_c(x_1^n) = \prod_{t=1}^{n} \frac{v(x_t/x_1^{t-1}) + \frac{1}{2}}{t - 1 + \frac{|A|}{2}},$$

KTSLRT algorithm

- $L_n(X^{\Delta}_{1:n})$ from combined use of Krichevsky-Trofimov estimator and the Arithmetic Encoder
  - Estimate source distribution via K-T estimator for a finite alphabet source as,

$$P_c(x_1^n) = \prod_{t=1}^{n} \frac{v(x_t/x_1^{t-1}) + \frac{1}{2}}{t - 1 + \frac{|A|}{2}},$$

  - Using AE with above estimate: $L_n(x_1^n) < -\log P_c(x_1^n) + 2$

KTSLRT algorithm

- $L_n(X^{\Delta}_{1:n})$ from combined use of Krichevsky-Trofimov estimator and the Arithmetic Encoder
  - Estimate source distribution via K-T estimator for a finite alphabet source as,

$$P_c(x_1^n) = \prod_{t=1}^{n} \frac{v(x_t/x_1^{t-1}) + \frac{1}{2}}{t - 1 + \frac{|A|}{2}},$$

  - Using AE with above estimate: $L_n(x_1^n) < -\log P_c(x_1^n) + 2$
- Universal codes defined via the K-T estimator and AE are nearly optimal

Performance Comparison LZLRT

$E_{DD} = 0.5\, E_{H_1}[N] + 0.5\, E_{H_0}[N]$ versus $P_E = 0.5\, P_{FA} + 0.5\, P_{MD}$

Performance Comparison LZLRT

$E_{DD} = 0.5\, E_{H_1}[N] + 0.5\, E_{H_0}[N]$ versus $P_E = 0.5\, P_{FA} + 0.5\, P_{MD}$

E_{DD}	$P_E = 0.05$	$P_E = 0.01$	$P_E = 0.005$
SPRT	3.21	4.59	6.29
GLR-Lai	5.0	8.53	12.83
LZSLRT	12.95	15.19	19.29

Table: $f_0 \sim \mathcal{N}(0,5)$ and $f_0 \sim \mathcal{N}(3,5)$

Performance Comparison LZLRT

$E_{DD} = 0.5\,E_{H_1}[N] + 0.5\,E_{H_0}[N]$ versus $P_E = 0.5\,P_{FA} + 0.5\,P_{MD}$

E_{DD}	$P_E = 0.05$	$P_E = 0.01$	$P_E = 0.005$
SPRT	3.21	4.59	6.29
GLR-Lai	5.0	8.53	12.83
LZSLRT	12.95	15.19	19.29

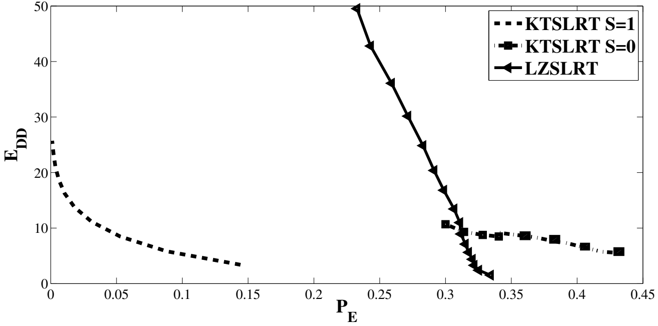
Table: $f_0 \sim \mathcal{N}(0,5)$ and $f_0 \sim \mathcal{N}(3,5)$

E_{DD}	$P_E = 0.05$	$P_E = 0.01$	$P_E = 0.005$
SPRT	7.45	10.86	18.23
GLR-Lai	18.21	29.65	33.42
LZSLRT	16.96	28.31	31.48

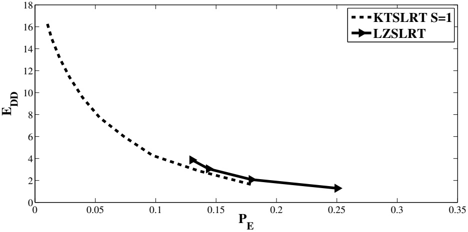
Table: $f_0 \sim \mathcal{P}(10,2)$ and $f_0 \sim \mathcal{P}(3,2)$

.

$f_1 \sim \mathcal{N}(0,5)$ and $f_0 \sim \mathcal{N}(0,1)$, 8 bit uniform quantizer.



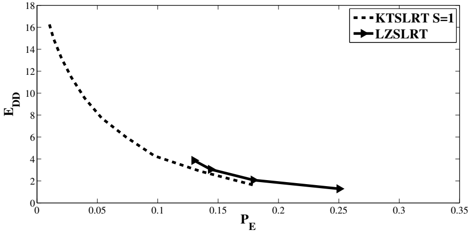
$f_1 \sim ln\mathcal{N}(3,3)$ and
$f_0 \sim ln\mathcal{N}(0,3)$



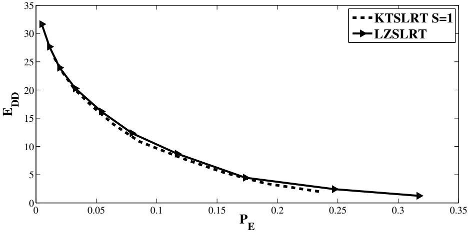
Performance Comparison

Lognormal Distribution and Pareto Distribution examples

$f_1 \sim ln\mathcal{N}(3,3)$ and $f_0 \sim ln\mathcal{N}(0,3)$



$f_1 \sim \mathcal{P}(3,2)$ and $f_0 \sim \mathcal{P}(10,2)$, support set $(2,10)$



Performance Comparison

Gaussian case ($f_1 \sim \mathcal{N}(0,5)$ and $f_0 \sim \mathcal{N}(0,1)$) with different estimators

1-NN differential entropy estimator is, $\gamma =$ Euler-Mascheroni constant

$$\hat{h}_n = \frac{1}{n}\sum_{i=1}^{n} \log \rho(i) + \log(n-1) + \gamma + 1,\ \rho(i) \triangleq \min_{j:1 \le j \le n, j \ne i} ||X_i - X_j||$$

Performance Comparison

Gaussian case ($f_1 \sim \mathcal{N}(0,5)$ and $f_0 \sim \mathcal{N}(0,1)$) with different estimators

1-NN differential entropy estimator is, $\gamma =$ Euler-Mascheroni constant

$$\hat{h}_n = \frac{1}{n}\sum_{i=1}^{n} \log \rho(i) + \log(n-1) + \gamma + 1,\ \rho(i) \triangleq \min_{j:1\leq j\leq n, j\neq i} ||X_i - X_j||$$

Kernel density estimator at a point z is

$$\hat{f}_n(z) = \frac{1}{w_n}\sum_{i=1}^{n} K\left(\frac{z - X_i}{w_n}\right),\ k\text{: kernel and } w_n\text{: bandwidth}$$

Performance Comparison

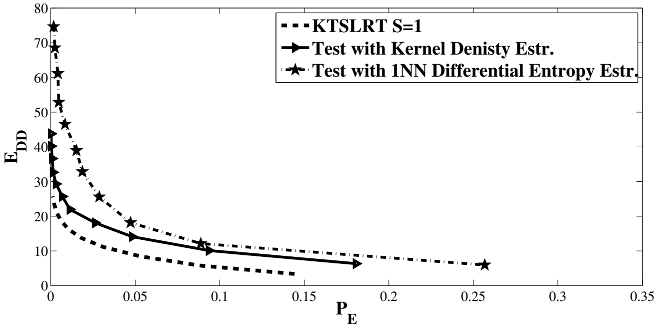
Gaussian case ($f_1 \sim \mathcal{N}(0,5)$ and $f_0 \sim \mathcal{N}(0,1)$) with different estimators

1-NN differential entropy estimator is, $\gamma =$ Euler-Mascheroni constant

$$\hat{h}_n = \frac{1}{n}\sum_{i=1}^{n} \log \rho(i) + \log(n-1) + \gamma + 1,\ \rho(i) \triangleq \min_{j:1\leq j\leq n, j\neq i} ||X_i - X_j||$$

Kernel density estimator at a point z is

$$\hat{f}_n(z) = \frac{1}{w_n}\sum_{i=1}^{n} K\left(\frac{z - X_i}{w_n}\right),\ k\text{: kernel and } w_n\text{: bandwidth}$$



Performance Comparison

Comparison with Hoeffding Test: $P_0 \sim B(8, 0.2)$, $P_1 \sim B(8, 0.5)$

Performance Comparison

Comparison with Hoeffding Test: $P_0 \sim B(8, 0.2)$, $P_1 \sim B(8, 0.5)$

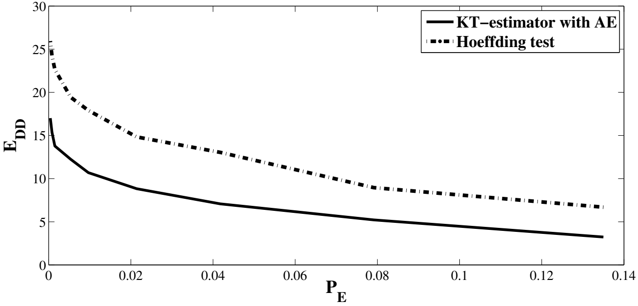
- Hoeffding test: Asymptotically optimal universal fixed sample size test for finite alphabet sources

Performance Comparison

Comparison with Hoeffding Test: $P_0 \sim B(8, 0.2)$, $P_1 \sim B(8, 0.5)$

- Hoeffding test: Asymptotically optimal universal fixed sample size test for finite alphabet sources
- $\delta_{FSS} = \mathbb{I}\{D(\Gamma^n || P_0) \geq \eta\}$

- Hoeffding test: Asymptotically optimal universal fixed sample size test for finite alphabet sources
- $\delta_{FSS} = \mathbb{I}\{D(\Gamma^n || P_0) \geq \eta\}$



- Sequential+Universal+Cooperative

- Sequential+Universal+Cooperative
- observations are i.i.d. at any CR and independent across CRs.

- Sequential+Universal+Cooperative
- observations are i.i.d. at any CR and independent across CRs.
- Algorithm: at SU's: , at FC:

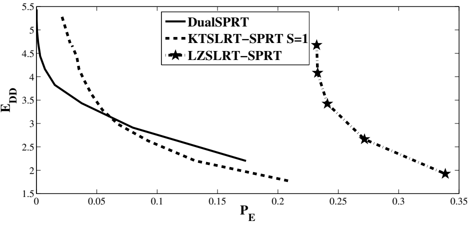
- Sequential+Universal+Cooperative
- observations are i.i.d. at any CR and independent across CRs.
- Algorithm: at SU's: LZSLRT/KTSLRT, at FC: SPRT

- Sequential+Universal+Cooperative
- observations are i.i.d. at any CR and independent across CRs.
- Algorithm: at SU's: LZSLRT/KTSLRT, at FC: SPRT
- noisy MAC between SUs and FC
- FC SPRT: binary hypothesis testing of g_{μ_1} vs $g_{-\mu_0}$

$b_1 = 1$, $b_0 = -1$, $I = 2$, $L = 5$ and $Z_k \sim \mathcal{N}(0,1)$

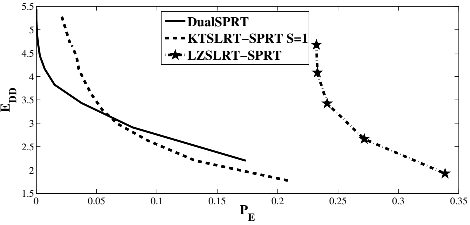
Decentralized case
Performance Comparison

$b_1 = 1,\ b_0 = -1,\ I = 2,\ L = 5$ and $Z_k \sim \mathcal{N}(0,1)$

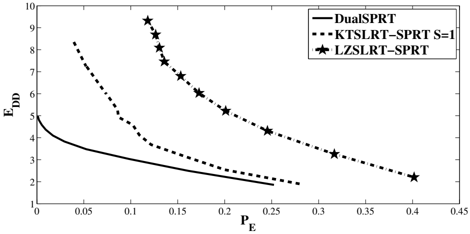


$f_{0,l} \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(0,5)$, for $1 \leq l \leq L$.

$b_1 = 1,\ b_0 = -1,\ l = 2,\ L = 5$ and $Z_k \sim \mathcal{N}(0, 1)$



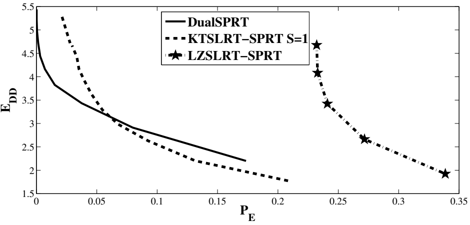
$f_{0,l} \sim \mathcal{N}(0, 1)$ and $f_{1,l} \sim \mathcal{N}(0, 5)$, for $1 \leq l \leq L$.



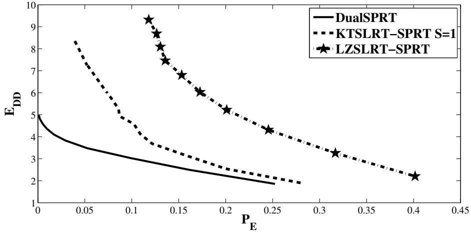
$f_{0,l} \sim \mathcal{P}(10, 2)$ and $f_{1,l} \sim \mathcal{P}(3, 2)$, for $1 \leq l \leq L$.

Decentralized case
Performance Comparison

$b_1 = 1,\ b_0 = -1,\ I = 2,\ L = 5$ and $Z_k \sim \mathcal{N}(0,1)$



$f_{0,l} \sim \mathcal{N}(0,1)$ and $f_{1,l} \sim \mathcal{N}(0,5)$, for $1 \leq l \leq L$.



$f_{0,l} \sim \mathcal{P}(10,2)$ and $f_{1,l} \sim \mathcal{P}(3,2)$, for $1 \leq l \leq L$.

Analysis using Perturbed Random Walk theory

$$\widehat{W_n} = S_n + \xi_n, \xi_n/n \to 0 \text{ a.s.}$$

Multihypothesis Decentralized Sequential Testing

(distributions completely known)

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis ($M > 2$), $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis ($M > 2$), $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i\left[\log \frac{f^k(X_1)}{f^j(X_1)}\right] =$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis $(M > 2)$, $H_i : X_n \sim f^i$, $i = 0, \dots, M-1$

$$E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = D(f^i || f^j) - D(f^i || f^k)$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis $(M > 2)$, $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i\left[\log \frac{f^k(X_1)}{f^j(X_1)}\right] = D(f^i||f^j) - D(f^i||f^k)$$

$$\min_{j \neq k} E_i\left[\log \frac{f^k(X_1)}{f^j(X_1)}\right] =$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis $(M > 2)$, $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = D(f^i || f^j) - D(f^i || f^k)$$

$$\min_{j \neq k} E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = \begin{cases} < 0 & \text{when } k \neq i \\ & \end{cases}$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis ($M > 2$), $H_i : X_n \sim f^i$, $i = 0, \dots, M-1$

$$E_i\left[\log \frac{f^k(X_1)}{f^j(X_1)}\right] = D(f^i||f^j) - D(f^i||f^k)$$

$$\min_{j \neq k} E_i\left[\log \frac{f^k(X_1)}{f^j(X_1)}\right] = \begin{cases} < 0 & \text{when } k \neq i \\ > 0 & \text{when } k = i \end{cases}$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis ($M > 2$), $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = D(f^i || f^j) - D(f^i || f^k)$$

$$\min_{j \neq k} E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = \begin{cases} < 0 & \text{when } k \neq i \\ > 0 & \text{when } k = i \end{cases}$$

- At local node l, $\{W_{n,l}^{k,j}, 0 \leq k, j \leq M-1\}$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis $(M > 2)$, $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = D(f^i || f^j) - D(f^i || f^k)$$

$$\min_{j \neq k} E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = \begin{cases} < 0 & \text{when } k \neq i \\ > 0 & \text{when } k = i \end{cases}$$

- At local node l, $\{W_{n,l}^{k,j}, 0 \leq k, j \leq M-1\}$

$$W_{n,l}^{k,j} = W_{n-1,l}^{k,j} + \log \frac{f_l^k(X_{n,l})}{f_l^j(X_{n,l})}, W_{0,l}^{k,j} = 0$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis $(M > 2)$, $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = D(f^i || f^j) - D(f^i || f^k)$$

$$\min_{j \neq k} E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = \begin{cases} < 0 & \text{when } k \neq i \\ > 0 & \text{when } k = i \end{cases}$$

- At local node l, $\{W_{n,l}^{k,j}, 0 \leq k, j \leq M-1\}$

$$\begin{aligned} W_{n,l}^{k,j} &= W_{n-1,l}^{k,j} + \log \frac{f_l^k(X_{n,l})}{f_l^j(X_{n,l})}, W_{0,l}^{k,j} = 0 \\ N_l &= \inf\{n : W_{n,l}^{k,j} > A \text{ for all } j \neq k \text{ and some } k\} \end{aligned}$$

Multihypothesis Decentralized Sequential Testing: Algorithm-1 (DMSLRT-1)

- M Hypothesis ($M > 2$), $H_i : X_n \sim f^i$, $i = 0, \ldots, M-1$

$$E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = D(f^i || f^j) - D(f^i || f^k)$$

$$\min_{j \neq k} E_i \left[\log \frac{f^k(X_1)}{f^j(X_1)} \right] = \begin{cases} < 0 & \text{when } k \neq i \\ > 0 & \text{when } k = i \end{cases}$$

- At local node l, $\{W_{n,l}^{k,j}, 0 \leq k, j \leq M-1\}$

$$\begin{aligned} W_{n,l}^{k,j} &= W_{n-1,l}^{k,j} + \log \frac{f_l^k(X_{n,l})}{f_l^j(X_{n,l})}, W_{0,l}^{k,j} = 0 \\ N_l &= \inf\{n : W_{n,l}^{k,j} > A \text{ for all } j \neq k \text{ and some } k\} \\ \text{or } N_l &= \inf\{n : \max_k \min_{j \neq k} W_{n,l}^{k,j} > A\} \end{aligned}$$

Algorithm-1 (DMSLRT-1) Contd.

- Use reflected random walks at local nodes, $\max(W_{n,l}, 0)$

Algorithm-1 (DMSLRT-1) Contd.

- Use reflected random walks at local nodes, $\max(W_{n,l}, 0)$
  Decision by node $l = H_i$
- At time $k \geq N_l$, node l transmits b_i

Algorithm-1 (DMSLRT-1) Contd.

- Use reflected random walks at local nodes, $\max(W_{n,l}, 0)$
  Decision by node $l = H_i$
- At time $k \geq N_l$, node l transmits b_i
- Instead of physical layer fusion, TDMA used.

Algorithm-1 (DMSLRT-1) Contd.

- Use reflected random walks at local nodes, $\max(W_{n,l}, 0)$
  Decision by node $l = H_i$
- At time $k \geq N_l$, node l transmits b_i
- Instead of physical layer fusion, TDMA used.
- FC uses the same test with hypotheses
  $G_m : Y_k \sim f_{FC}^m = \mathcal{N}(b_m, \sigma^2)$

Numerical results

$H_m : X_{k,l} \sim \mathcal{N}(m, 1),\ m{=}0,\ldots,\ 4,\ \text{No of local nodes}{=}5$

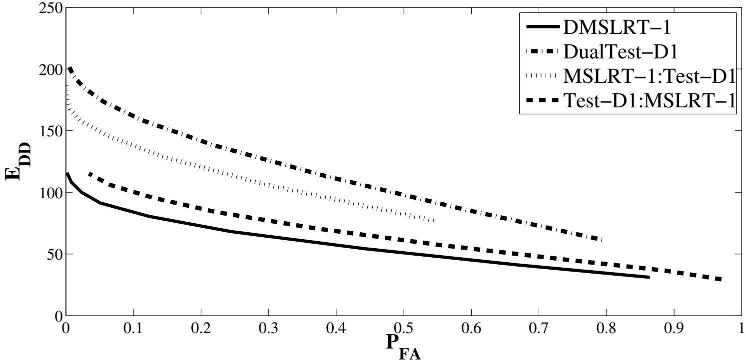


Figure: Comparison among different Multihypothesis schemes

Algorithm-2 (DMSLRT-2)

- Reduce false alarms caused by Gaussian noise before first transmission from local nodes.

Algorithm-2 (DMSLRT-2)

- Reduce false alarms caused by Gaussian noise before first transmission from local nodes.
- FC statistic is modified as

$$F_n^{i,j} = \widehat{F}_n^{i,0} - \widehat{F}_n^{j,0}, \text{ where } \widehat{F}_n^{i,0} = \max\left(\widehat{F}_{n-1}^{i,0} + \log \frac{f_{FC}^i(Y_n)}{f_Z(Y_n)}, 0\right).$$

Algorithm-2 (DMSLRT-2)

- Reduce false alarms caused by Gaussian noise before first transmission from local nodes.
- FC statistic is modified as

$$F_n^{i,j} = \widehat{F}_n^{i,0} - \widehat{F}_n^{j,0}, \text{ where } \widehat{F}_n^{i,0} = \max\left(\widehat{F}_{n-1}^{i,0} + \log\frac{f_{FC}^{i}(Y_n)}{f_Z(Y_n)}, 0\right).$$

- Expected drift of $\widehat{F}_n^{i,0} > 0$ only when $E[Y_n] > b_i/2$

Algorithm-2 (DMSLRT-2)

- Reduce false alarms caused by Gaussian noise before first transmission from local nodes.
- FC statistic is modified as

$$F_n^{i,j} = \widehat{F}_n^{i,0} - \widehat{F}_n^{j,0}, \text{ where } \widehat{F}_n^{i,0} = \max\left(\widehat{F}_{n-1}^{i,0} + \log\frac{f_{FC}^i(Y_n)}{f_Z(Y_n)}, 0\right).$$

- Expected drift of $\widehat{F}_n^{i,0} > 0$ only when $E[Y_n] > b_i/2$
- Positive b_i's make $E[F_n^{i,j}]$ negligible before first transmission

Numerical results

SNRs (-10 dB, -6 dB, 0 dB and 6 dB), PU with SNR -10 dB uses the channel.

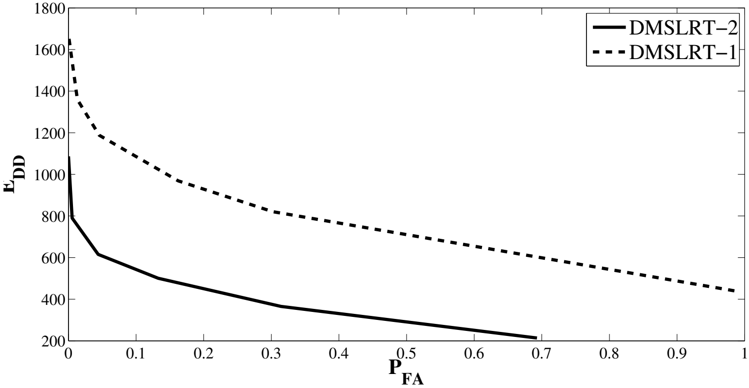


Figure: Comparison between MDSLRT-1 and MDSLRT-2

Analysis of Algorithm-1(DMSLRT-1)

- E_{DD} **Analysis at SU**: Dominant event-Reflected random walk with minimum positive expected drift

Analysis of Algorithm-1(DMSLRT-1)

- E_{DD} **Analysis at SU**: Dominant event-Reflected random walk with minimum positive expected drift

$$E_{DD}^{l} = E_i[N_l] \approx \frac{A}{\min_{j \neq i} D(f_l^i || f_l^j)}$$

Analysis of Algorithm-1(DMSLRT-1)

- E_{DD} **Analysis at SU**: Dominant event-Reflected random walk with minimum positive expected drift

$$E_{DD}^{l} = E_i[N_l] \approx \frac{A}{\min_{j \neq i} D(f_l^i || f_l^j)}$$

- Nonlinear renewal theory to take care of overshoots

Analysis of Algorithm-1(DMSLRT-1)

- E_{DD} **Analysis at SU**: Dominant event-Reflected random walk with minimum positive expected drift

$$E_{DD}^{l} = E_i[N_l] \approx \frac{A}{\min_{j \neq i} D(f_l^i || f_l^j)}$$

- Nonlinear renewal theory to take care of overshoots
- P_{FA} **Analysis at SU**: Dominant event in $\{W_{n,l}^{k,j}, \ k \neq i\}$-when the expected drift is most negative.

Analysis of Algorithm-1(DMSLRT-1)

- E_{DD} **Analysis at SU**: Dominant event-Reflected random walk with minimum positive expected drift

$$E_{DD}^{l} = E_i[N_l] \approx \frac{A}{\min_{j \neq i} D(f_l^i || f_l^j)}$$

- Nonlinear renewal theory to take care of overshoots
- P_{FA} **Analysis at SU**: Dominant event in $\{W_{n,l}^{k,j},\, k \neq i\}$-when the expected drift is most negative.
- $N_l^{k,j} = \inf\{n : W_{n,l}^{k,j} > A\}$; $P_{FA}^{l} \approx P(\min_{k \neq i} N_l^{k,i} < N_l^{i,\widehat{j}})$
$\widehat{j} = \operatorname{argmin}_{j \neq i} D(f_l^i || f_l^j)$

Analysis of Algorithm-1(DMSLRT-1)

- E_{DD} **Analysis at SU**: Dominant event-Reflected random walk with minimum positive expected drift

$$E_{DD}^{l} = E_i[N_l] \approx \frac{A}{\min_{j\neq i} D(f_l^i || f_l^j)}$$

- Nonlinear renewal theory to take care of overshoots
- P_{FA} **Analysis at SU**: Dominant event in $\{W_{n,l}^{k,j},\ k \neq i\}$-when the expected drift is most negative.
- $N_l^{k,j} = \inf\{n : W_{n,l}^{k,j} > A\}$; $P_{FA}^{l} \approx P(\min_{k\neq i} N_l^{k,i} < N_l^{i,\widehat{j}})$
  $\widehat{j} = \text{argmin}_{j\neq i} D(f_l^i || f_l^j)$
- E_{DD} **Analysis of DMSLRT-1**: $E_{DD} \approx E_i[\max_l N_l] + E_i[N_{FC}]$

Comparisons of Analysis with simulations

Threshold (A)	E_{DD} Simln.	E_{DD} Anal
100	157.54	141.61
120	186.98	169.34
140	216.31	197.07

Table: At SU: Comparison of E_{DD} obtained via simulation and analysis.

Comparisons of Analysis with simulations

Threshold (A)	E_{DD} Simln.	E_{DD} Anal
100	157.54	141.61
120	186.98	169.34
140	216.31	197.07

Table: At SU: Comparison of E_{DD} obtained via simulation and analysis.

Threshold (A)	P^{l}_{FA} Simln.	P^{l}_{FA} Anal
8	0.0138	0.0296
10	0.0043	0.0093
20	$5.00E-5$	$1.81E-5$

Table: At SU: Comparison of P_{FA} obtained via simulation and analysis.

Comparisons of Analysis with simulations

Threshold (A)	E_{DD} Simln.	E_{DD} Anal
100	157.54	141.61
120	186.98	169.34
140	216.31	197.07

Table: At SU: Comparison of E_{DD} obtained via simulation and analysis.

Threshold (A)	P^l_{FA} Simln.	P^l_{FA} Anal
8	0.0138	0.0296
10	0.0043	0.0093
20	$5.00E-5$	$1.81E-5$

Table: At SU: Comparison of P_{FA} obtained via simulation and analysis.

A	B	E_{DD} Simln.	E_{DD} Anal
10	80	116.79	133.63
10	90	144.04	147.49
10	100	163.54	161.36

Table: DMSLRT-1: Comparison of E_{DD} obtained via simulation and analysis.

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework
- First asymptotically optimal decentralized sequential hypothesis test for noisy MAC; Modified to perform well in non-asymptotic way; performance analysis, numerical comparisons

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework
- First asymptotically optimal decentralized sequential hypothesis test for noisy MAC; Modified to perform well in non-asymptotic way; performance analysis, numerical comparisons
- Parameter uncertainty case.

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework
- First asymptotically optimal decentralized sequential hypothesis test for noisy MAC; Modified to perform well in non-asymptotic way; performance analysis, numerical comparisons
- Parameter uncertainty case.
- Universal sequential spectrum sensing algorithms-unknown channel statistic scenario.

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework
- First asymptotically optimal decentralized sequential hypothesis test for noisy MAC; Modified to perform well in non-asymptotic way; performance analysis, numerical comparisons
- Parameter uncertainty case.
- Universal sequential spectrum sensing algorithms-unknown channel statistic scenario.
- Universal sequential tests: discrete alphabet source

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework
- First asymptotically optimal decentralized sequential hypothesis test for noisy MAC; Modified to perform well in non-asymptotic way; performance analysis, numerical comparisons
- Parameter uncertainty case.
- Universal sequential spectrum sensing algorithms-unknown channel statistic scenario.
- Universal sequential tests: discrete alphabet source (asymptotic properties derived), continuous alphabet source, decentralized scenario

Conclusions

- Cooperating Spectrum Sensing algorithms in sequential detection framework
- First asymptotically optimal decentralized sequential hypothesis test for noisy MAC; Modified to perform well in non-asymptotic way; performance analysis, numerical comparisons
- Parameter uncertainty case.
- Universal sequential spectrum sensing algorithms-unknown channel statistic scenario.
- Universal sequential tests: discrete alphabet source (asymptotic properties derived), continuous alphabet source, decentralized scenario
- Decentralized multihypothesis Sequential tests: performance analysis, numerical comparisons

Thank You

Publications based on this Thesis

1. J. K. Sreedharan and V. Sharma, "Spectrum sensing using distributed sequential detection via noisy MAC", submitted to journal.
2. J. K. Sreedharan and V. Sharma. "Nonparametric distributed sequential detection via universal source coding", in *Information Theory and Applications Workshop (ITA)*, California, USA, Feb 2013.
3. J. K. Sreedharan and V. Sharma, "Spectrum Sensing via Universal Source Coding", in *Proc. IEEE Global Communications Conference (GLOBECOM)*, California, USA, Dec 2012.
4. K. S. Jithin and V. Sharma, "Novel algorithms for distributed sequential hypothesis testing", in *Proc. 49th Annual Allerton Conference on Communication, Control and Computing*, Illinois, USA, Sep 2011 (invited paper).
5. J. K. Sreedharan and V. Sharma, "A novel algorithm for cooperative distributed sequential spectrum sensing in Cognitive Radio", in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Cancun, Mexico, Mar 2011.
6. K. S. Jithin, V. Sharma, and R. Gopalarathnam, "Cooperative distributed sequential spectrum sensing", in *Proc. IEEE National Conference on Communication (NCC)*, Bangalore, India, Jan 2011.

Lemmas to support DualSPRT analysis

Lemma 1

For $i = 0, 1$, $P_i(N_l = N_l^i) \to 1$ as $\gamma \to \infty$ and $P_i(N = N^i) \to 1$ as $\gamma \to \infty$ and $\beta \to \infty$.

Lemma 2

Under H_i, $i = 0, 1$ and $j \neq i$,

(a) $|N_l - N_l^i| \to 0$ a.s. as $\gamma \to \infty$ and $\lim_{\gamma \to \infty} \frac{N_l}{\gamma} = \lim_{\gamma \to \infty} \frac{N_l^i}{\gamma} = \frac{1}{D(f_{i,l} || f_{j,l})}$ a.s. and in L^1.

(b) $|N - N^i| \to 0$ a.s. and $\lim \frac{N}{\beta} = \lim \frac{N^i}{\beta}$ a.s. and in L^1, as $\gamma \to \infty$ and $\beta \to \infty$.

Lemma 3

Let t_k be the time when k local nodes have made the decision. Under H_i, $i = 0, 1$, as $\gamma \to \infty$,
$P_i(\text{Decision at time } t_k \text{ is } H_i \text{ and } t_k \text{ is the } k^{th} \text{ order statistics of } N_1^i, N_2^i, \ldots, N_L^i) \to 1$.

CR block diagram

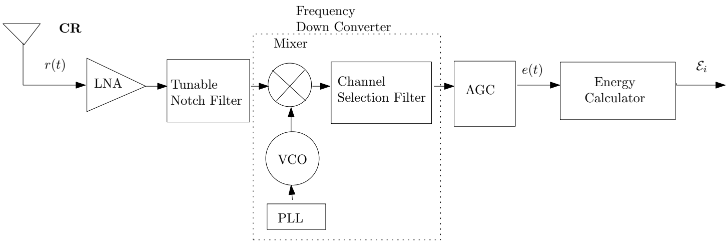


Figure: Block diagram of the receiver implementation at a CR

$P_{H_1}(FA\ before\ t_1)$

$$
\begin{aligned}
&= \sum_{k=1}^{\infty} P\Big[\{F_k < -\theta\} \cap_{n=1}^{k-1} \{F_n > -\theta\}\Big| t_1 > k\Big] P[t_1 > k] \\
&= \sum_{k=1}^{\infty} \Big(P[F_k < -\theta| \cap_{n=1}^{k-1} \{F_n > -\theta\}]\ P[\cap_{n=1}^{k-1}\{F_n > -\theta\}]\Big) \\
&\quad P[t_1 > k] \\
&\overset{(A)}{=} \sum_{k=1}^{\infty} \Big(P[F_k < -\theta|F_{k-1} > -\theta]\ P[\inf_{1\le n\le k-1} F_n > -\theta]\Big) \\
&\quad P[t_1 > k] \\
&\overset{(B)}{\geq} \sum_{k=1}^{\infty} \Big(\int_{c=0}^{2\theta} P[S_k < -c] f_{F_{k-1}}\{-\theta + c\} dc\Big) \\
&\quad \Big(1 - 2P[F_{k-1} < -\theta]\Big)\Big(\prod_{l=1}^{L}(1 - \Phi_{\tau_{\gamma,l}}(k)\Big)
\end{aligned}
$$

(A) is because of the Markov property of the random walk. (B) is due to the inequality, $P[\sup_{k\le n} F_k \geq \theta] \leq 2P[F_n \geq \theta]$ for the Gaussian R.W

Gaussian Mean change approximation of Energy Detector

- $X_{k,l}$ are a summation of energy of N samples received by the l^{th} Cognitive Radio
- For large N, the pre and post change distributions of $X_{k,l}$ can be approximated by Gaussian distributions: $f_{0,l} \sim \mathcal{N}(\sigma^2, 2\sigma^4/N)$ and $f_{1,l} \sim \mathcal{N}(P_l + \sigma^2, 2(P_l + \sigma^2)^2/N)$ where P_l is the received power at the l^{th} CR node and noise $Z_{k,l} \sim \mathcal{N}(0, \sigma^2)$.
- Under low SNR conditions $(P_l + \sigma^2)^2 \approx \sigma^4$ and hence $X_{k,l}$ are Gaussian distributed with mean change under H_0 and H_1
- Take $X_{k,l} - \sigma^2$ as the data for l^{th} node SPRT.

Gaussian Mean change under log normal modelling

- $X_{k,l}$ is the received power in decibels.
- pre change distribution of $X_{k,l}$ is $\mathcal{N}(\mu_0, {\sigma_l}^2)$ and post change distribution $\mathcal{N}(\mu_0 + P_l, {\sigma_l}^2)$
- Secondary nodes are using Energy detector
- Under H_0 the uncertainty in noise and interference power is assumed to be log normally distributed. i.e log of the received power is Gaussian. μ_0 is mean noise power. ${\sigma_l}^2$ is the uncertainty in noise power.
- Under H_1, P_l- mean increase in received power due to the presence of primary. $P_l = 10 \log_{10}(1 + SNR)$ dB.
- Log normal distribution is valid under H_1 as its used for modelling shadowing.

Nonlinear renewal theory argument for MSLRT-1 analysis

Under mild conditions the limiting distribution of the excess of a random walk over a fixed threshold does not change by the addition of a slowly changing nonlinear term.

$$E_i[N_l] \approx \frac{A + \mathcal{X}_l^{i,\widehat{j}} + \mathcal{B}_l^{i,\widehat{j}}}{D(f_l^i || f_l^{\widehat{j}})},$$

where $\mathcal{X}_l^{i,\widehat{j}} = \frac{E_i[(R_{1,l}^{i,\widehat{j}})^2]}{2E_i[(R_{1,l}^{i,\widehat{j}})^2]} - \sum_{n=1}^{\infty} n^{-1} E_i S_{n,l}^{-\,i,\widehat{j}}$ and $\mathcal{B}_l^{i,\widehat{j}} = -\sum_{n=1}^{\infty} n^{-1} E_i S_{n,l}^{-\,i,\widehat{j}}$

$$R_{k,l}^{i,\widehat{j}} = \log(f_l^i(X_{k,l}/f_l^{\widehat{j}}(X_{k,l}) \text{ and } S_{n,l}^{-\,i,\widehat{j}} = -\min(0, \sum_{k=1}^{n} R_{k,l}^{i,\widehat{j}}).$$