A Step Towards the Cognitive Radar: Target Detection under Nonstationary Clutter

Murat Akcakaya

Department of Electrical and Computer Engineering

University of Pittsburgh

Email: akcakaya@pitt.edu

Satyabrata Sen[*], Yijian Xiang[†], and Arye Nehorai[†]

Oak Ridge National Laboratory[*]

Washington University in St. Louis[†]

International Conference on InfoSymbiotics/DDDAS
August 7-9, 2017

Acknowledgement

This material is based upon the work supported by the Air Force Office of Scientific Research (AFOSR), the DDDAS Program, under Grant No. FA9550-16-1-0386. We thank AFOSR/DDDAS for supporting this project.

Outline

- Introduction
  - Challenges
  - Motivations
  - Our approach
  - Relevance to DDDAS
- Target Detection Method
  - Change point detection
  - CUSUM
  - Extended CUSUM
- Numerical Results
- Conclusions and Future Work

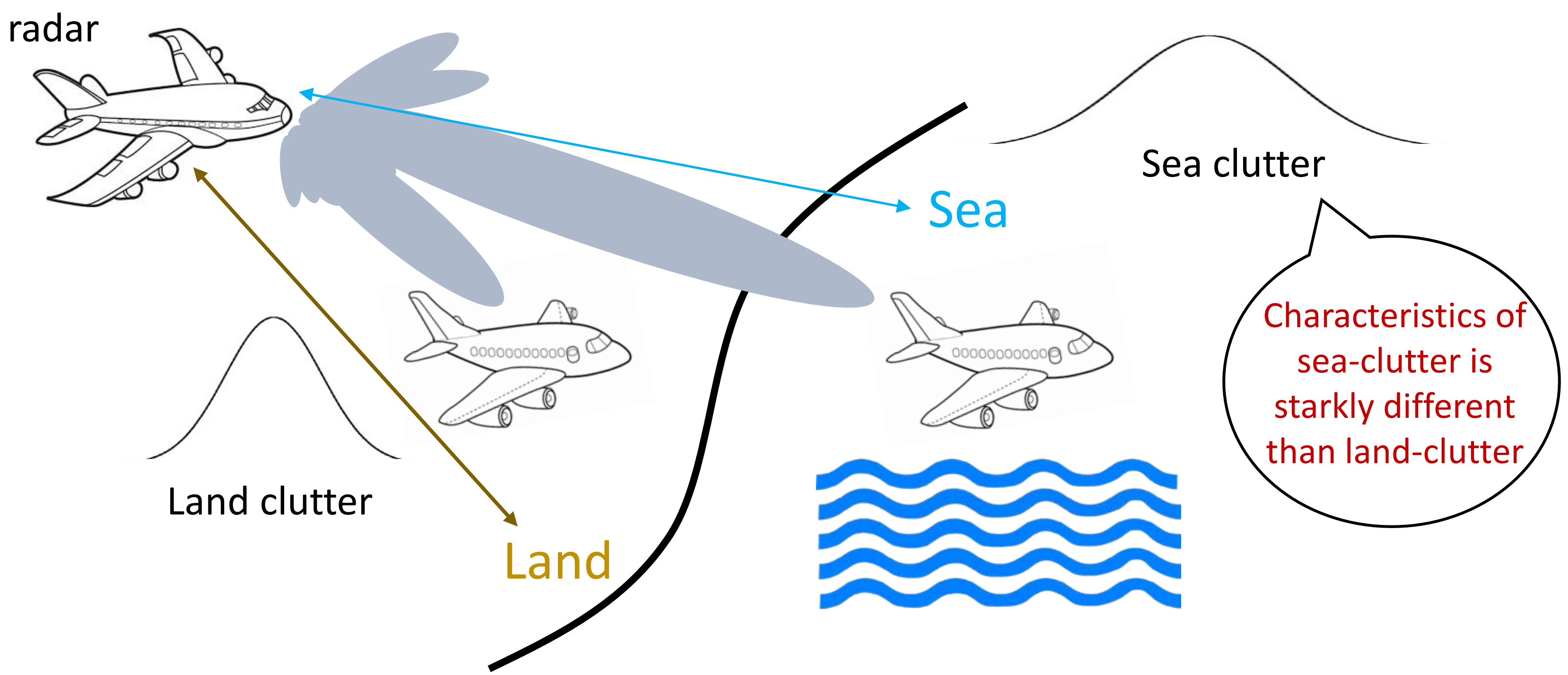
Introduction

Problem Description

- Goal:
  - Improve target detection capability in the presence of nonstationary clutter.

- Challenges:
  - Characteristics of the clutter can vary enormously depending on radar-operational scenarios (e.g., weather change, dynamic target, and changes in the environment).
  - However, most existing radar algorithms
    - assume stationary, parametric distributions for the clutter and noise processes,
    - consider the availability of a dictionary of possible probability density functions (pdfs),
    - do not address the problem of estimating the time instant when the characteristics of the clutter changes.
  - Therefore, the radar performance can severely deteriorate when the inherent nonstationary characteristics are not accounted for.

Problem Description (cont.)



An illustrative example of target detection in nonstationary environment

Motivation

- Cognitive radar framework conceptualizes an advanced radar system that
  - senses its scenario effectively,
  - learns from its experience,
  - adapts to the changes.

- Thus, in a nonstationary environment, a cognitive radar provides a unifying approach by
  - autonomously learning and estimating the change-point in the statistical characteristics of the scenario,
  - incorporating the newly learned environment statistics into the target detection algorithm.

Our Approach

- Employ a data-driven active drift detection algorithm
  - to detect changes of statistical characteristics of the environment.

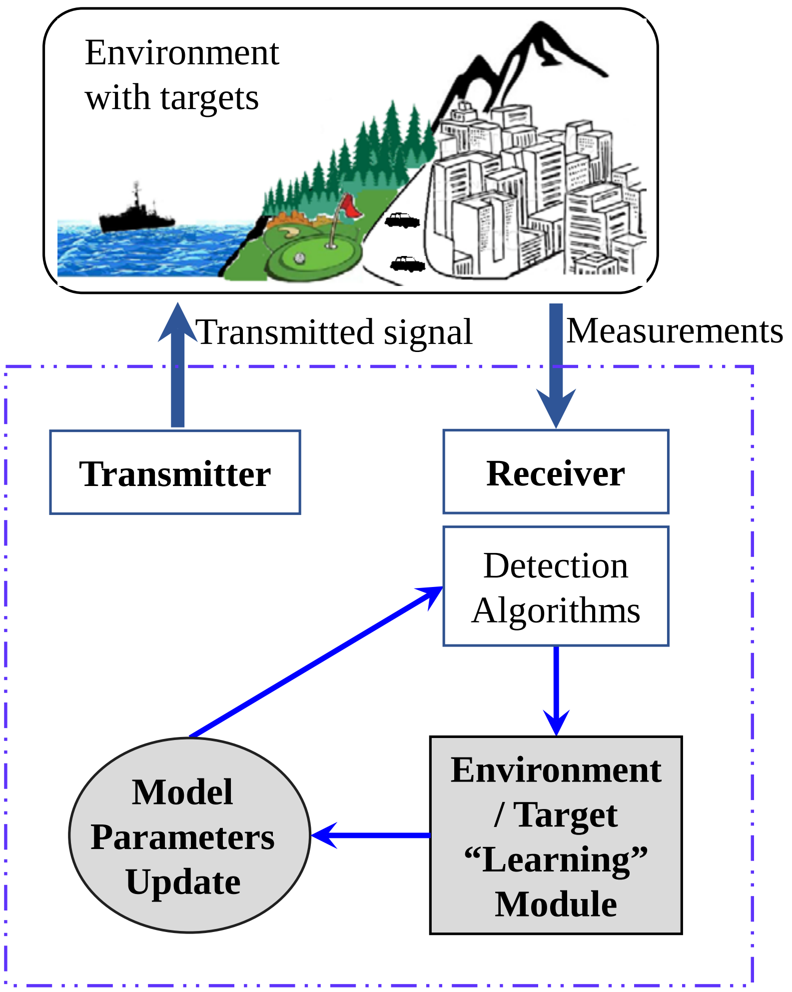
sensing

- Apply incremental learning (i.e., learning under concept drift) algorithms
  - to learn the new statistical characteristics of the environment.

learning

- Integrate the newly learned environment characteristics in the detection algorithm
  - to adapt to the changes in the environment.

adapting



Relevance to DDAS

- Our proposed technique with adaptive learning and cognitive augmentation is of high relevance
  - to remotely piloted radar systems
  - for substantially better intelligence, surveillance, and reconnaissance (ISR) capability

- It is also germane to the key science and technology frontiers of DDDAS:
  - Applications modeling →
    - Developing *data-driven* learning algorithms to detect any significant change in the underlying distribution of the model
    - Jointly combining newly measured data with the previously-stored calibration data to learn the environment model
  - Advances in Mathematical and Statistical Algorithms →
    - Incorporating learned model parameters into radar algorithms to provide robust and improved detection performance
  - Application Measurement Systems and Methods → [*on-going work, not part of this presentation*]
    - Synthesizing the next transmit waveforms based on the updated model parameters
    - Using the designed waveforms as *control-agents* to dynamically steer the next received measurements
  - Software Infrastructures and other systems software → [*future work, not part of this presentation*]
    - Addressing the near real-time implementation issues of the proposed signal processing algorithms.

Target Detection Method

Data-Driven Active Drift Detection

A typical change point detection problem:

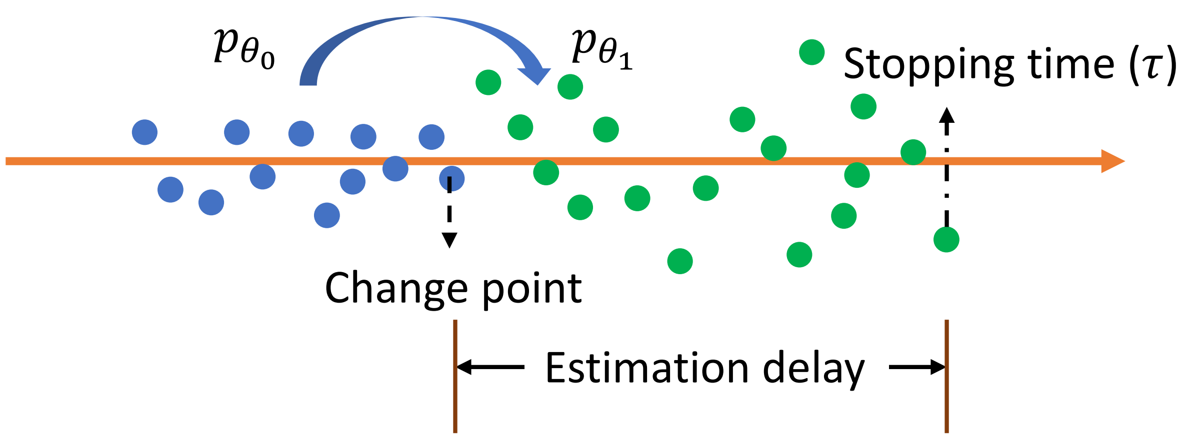
At current time N, decide from two hypotheses:

$$\mathcal{D}_0: \quad \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N \sim p_{\theta_0}$$

$$\mathcal{D}_1: \quad \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k_0} \sim p_{\theta_0}, \quad \boldsymbol{y}_{k_0+1}, \ldots, \boldsymbol{y}_N \sim p_{\theta_1}$$

where

- $\boldsymbol{y}_n$: observed data at time n;
- k_0: change point;
- $p_{\theta_0}, p_{\theta_1}$: two different distributions before and after the change point.



Data-Driven Active Drift Detection (cont.)

Change point detection formulation for radar detection problem:

$$\tau = \text{argmin}_{\tau} \sup_{n\geq 1} \text{ess sup } E_n[(\tau - n)^+ | \mathcal{Y}_0^c, \ldots, \mathcal{Y}_{n-1}^c] \quad (1)$$

$$\text{such that} \quad E_{\infty}[\tau] \geq \alpha$$

where

- $C_c,\ c \in \{0,1\}$: two classes representing the absence and presence of the target;
- $\mathcal{Y}_0 = \{\boldsymbol{y}_{0,t}, r_{0,t}\}, t = 1, \ldots, T_0$: calibration data set for two classes,
  - $\boldsymbol{y}_{0,t}$: t_{th} radar measurement
  - $r_{0,t} = 0$ if $\boldsymbol{y}_{0,t} \in C_0$ and $r_{0,t} = 1$ if $\boldsymbol{y}_{0,t} \in C_1$
  - T_0 : number of measurements in the calibration data;
- $\mathcal{Y}_n^c$: radar data observed at measurement time n from class c;
- τ : stopping time;
- $(x)^+ = \max(0, x)$;
- $E_n[\cdot]$: expectation with respect to p_n^c,
  - p_n^c : probability distribution when the change occurs at n;
- $E_{\infty}[\tau]$: mean time between false alarms, assuming that change never happens;
- ess sup : essential supremum of a set of random variables $\mathcal{X}$; it is a random variable Z such that (i) $P(Z \geq X) = 1\ \forall X \in \mathcal{X}$; and (ii) $P(Y \geq X) = 1, \forall X \in \mathcal{X} \rightarrow P(Y \geq Z) = 1, \forall X \in X$.

Data-Driven Active Drift Detection (cont.)

CUmulative SUM (CUSUM) algorithm:

– provides the optimum stopping time τ for the problem in (1)

$$\tau = \inf\left\{n \geq 1: \left(g(\mathcal{Y}_0^c, \ldots, \mathcal{Y}_n^c) = R_n - \min_{1 \leq k \leq n} R_k\right) \geq b\right\}$$

where

- $R_k = \sum_{i=1}^{k} \ln \frac{p_{\theta_1^c}(\mathcal{Y}_i^c)}{p_{\theta_0^c}(\mathcal{Y}_i^c)}$
- n : batch index
- b : a predefined threshold

Parameter training procedure:

i. Find an estimate of θ_0^c using
   a) calibration data $\mathcal{Y}_0^c$, $c \in \{0, 1\}$
   b) knowledge of the parametric distribution underlying the data

ii. Compute the confidence intervals for θ_0^c

iii. Assign the confidence interval extremas to θ_1^c

iv. Assign $b = \max_{1 \leq t \leq T_0} g(\mathcal{Y}_0^c)$, where $\mathcal{Y}_0^c = \boldsymbol{y}_{0,t} | r_{0,t} = c$, $t = 1, \ldots, T_0$

Data-Driven Active Drift Detection (cont.)

Extended CUSUM algorithm:

– when no prior knowledge is available about the distribution of $\mathcal{Y}_n^c$, $c \in \{0, 1\}$

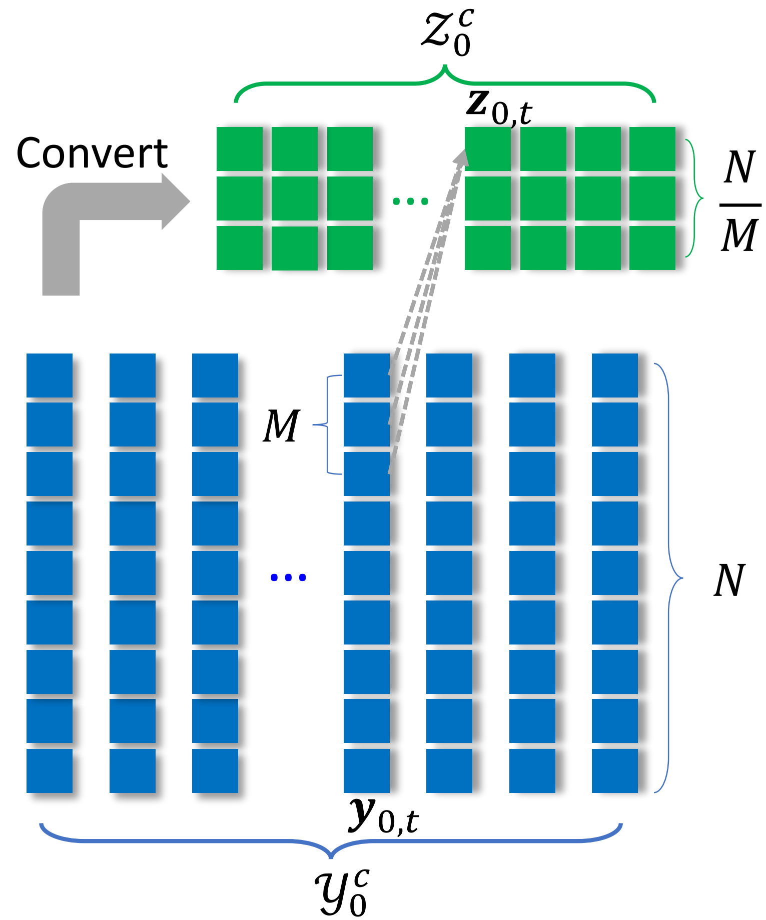
- Convert measurements into new samples (take $\mathcal{Y}_0^c$ as an example):

Convert

$$\mathcal{Y}_0^c = \{\boldsymbol{y}_{0,t} | r_t = c\}$$

$$\begin{cases} z_{0,t,i} = \dfrac{1}{M} \displaystyle\sum_{v=(i-1)M+1}^{iM} y_{0,t,v} \\ \boldsymbol{z}_{0,t} = \left[z_{0,t,1}, \dots, z_{0,t,N/M} \right]^T \\ \boxed{\mathcal{Z}_0^c = \{\boldsymbol{z}_{0,t} | r_t = c\}} \end{cases}$$

where $y_{0,t,v}$ and $z_{0,t,i}$ are v_{th} and i_{th} entries of vectors $\boldsymbol{y}_{0,t}$ and $\boldsymbol{z}_{0,t}$, respectively.



Data-Driven Active Drift Detection (cont.)

Extended CUSUM algorithm (cont.):

- Use new samples to employ CUSUM algorithm:

$$\tau = \inf\left\{n \geq 1: \left(g(\mathcal{Z}_0^c, \ldots, \mathcal{Z}_n^c) = R_n - \min_{1 \leq k \leq n} R_k\right) \geq b\right\}$$

where

- $R_k = \sum_{i=1}^{k} \ln \frac{p_{\theta_1^c}(\mathcal{Z}_i^c)}{p_{\theta_0^c}(\mathcal{Z}_i^c)}$
- n is the batch index
- b is a predefined threshold

Parameter training procedure:

i. Find an estimate of θ_0^c using the calibration data $\mathcal{Z}_0^c$, $c \in \{0, 1\}$

ii. Compute the confidence intervals for θ_0^c

iii. Assign the confidence interval extremas to θ_1^c

iv. Assign $b = \max_{1 \leq t \leq T_0} g(\mathcal{Z}_0^c)$, where $\mathcal{Z}_0^c = \mathbf{z}_{0,t} | r_{0,t} = c$, $t = 1, \ldots, T_0$

Numerical Results

Simulation Setup

Assumptions:

- Radar collects measurements
  - from multiple range cells (indexed by j)
  - over a sequence of coherent processing intervals (CPIs) (indexed by k)
- Radar receives and processes N temporal samples in each CPI
- Target is assumed to be present in the $j = 1$ range cell
- Target remains in a cell during the entire processing of $k = 1, 2, \dots, K$ CPIs
- Target response is considered to be known and constant (denoted as a)

Detection problem (of j_{th} range cell at k_{th} CPI):

$$\begin{cases} \mathcal{H}_0\colon\ \boldsymbol{y}_k^{(j)} = \boldsymbol{n}_k^{(j)} \\ \mathcal{H}_1\colon\ \boldsymbol{y}_k^{(j)} = a\mathbf{1} + \boldsymbol{n}_k^{(j)}, \end{cases} \quad \text{for } k = 1,2,\dots,K,\ j = 0,1,2\dots,$$

where each vector is of dimension $N \times 1$.

Nonstationary modeling:

- For $k = 1, 2, \dots, k_0$, we have $\boldsymbol{n}_k^{(j)} \sim p_{n,\theta_0}$, i.e., $\boldsymbol{y}_k^{(j)} \sim p_{\theta_0^c}$
- For $k = k_0 + 1, \dots, K$, we have $\boldsymbol{n}_k^{(j)} \sim p_{n,\theta_1}$, i.e., $\boldsymbol{y}_k^{(j)} \sim p_{\theta_1^c}$

Simulation Setup (cont.)

Parameter training procedure:

Suppose we have T_0^c training data for class c, i.e., $\mathcal{Y}_0^c = \{\boldsymbol{y}_{0,t}^c\}$ for $t = 1, \dots, T_0^c$.

- If the noise before change point is known as $\boldsymbol{n}_k^{(j)} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$, then

  1. $\theta_0^1 = [a, \sigma^2]$, $\theta_0^0 = [0, \sigma^2]$
  2. Compute sample mean $\hat{\mu}^c$ and sample variance $\hat{s}^{2c}$

$$\hat{\mu}^c = \frac{1}{T_0^c N} \sum_{t=1}^{T_0^c} \mathbf{1}^T \boldsymbol{y}_{0,t}^c \qquad \hat{s}^{2c} = \frac{1}{T_0^c N - 1} \sum_{t=1}^{T_0^c} \left(\boldsymbol{y}_{0,t}^c - \hat{\mu}^c \mathbf{1}\right)^T \left(\boldsymbol{y}_{0,t}^c - \hat{\mu}^c \mathbf{1}\right)$$

  3. Take the lower/upper limits on the 95% confidence intervals of mean and variance

$$a_{ll}^c = \hat{\mu}^c - 1.96 \frac{\sigma}{\sqrt{T_0^c N}} \qquad a_{ul}^c = \hat{\mu}^c - 1.96 \frac{\sigma}{\sqrt{T_0^c N}}$$

$$\sigma_{ll}^{2c} = \frac{(T_0^c N - 1)\hat{s}^{2c}}{\mathcal{X}_{T_0^c N-1}^2(\alpha/2)} \qquad \sigma_{ll}^{2c} = \frac{(T_0^c N - 1)\hat{s}^{2c}}{\mathcal{X}_{T_0^c N-1}^2(1-\alpha/2)} \qquad \text{where } \alpha = 0.05$$

  4. Evaluate $g_n = R_n - \min_{1 \le k \le n} R_k$, with $R_k = \sum_{t=1}^{k} \ln \frac{p_{\theta_1^c}(\boldsymbol{y}_{0,t}^c)}{p_{\theta_0^c}(\boldsymbol{y}_{0,t}^c)}$ having parameters from θ_0^c to either $\theta_1^c = [a_{ll}^c, \sigma_{ll}^{2c}]$, $\theta_1^c = [a_{ll}^c, \sigma_{ul}^{2c}]$, $\theta_1^c = [a_{ul}^c, \sigma_{ll}^{2c}]$, or $\theta_1^c = [a_{ul}^c, \sigma_{ul}^{2c}]$
  5. Let $b = \max_{1 \le n \le T_0^c} g_n$

Simulation Setup (cont.)

Parameter training procedure (cont.):

- If the noise $\boldsymbol{n}_k^{(j)}$ distribution is unknown, then the conversion for the extended CUSUM algorithm is used:
  - $\theta_0^c = [\hat{\mu}_{\boldsymbol{z}}^c, \hat{s}_{\boldsymbol{z}}^{2c}]$, where $\hat{\mu}_{\boldsymbol{z}}^c$ and $\hat{s}_{\boldsymbol{z}}^{2c}$ are sample mean and variance of $\mathcal{Z}_0^c = \{\boldsymbol{z}_{0,t}^c\}$ (training data after the conversion) under $\mathcal{H}_c$
  - θ_1^c and b can be trained similar to the above.

Change Point Detection:

Once θ_0^c, θ_1^c, and b are obtained, then CUSUM (or extended CUSUM) algorithm is used to detect the change point.

Adaptive Target Detection:

Once a change in the clutter distribution is detected, the proposed radar accordingly modifies the log-likelihood ratio

$$\underbrace{\ln\left[\frac{p_{\theta_0^c}\left(\boldsymbol{y}_k^{(j)}|\mathcal{H}_1\right)}{p_{\theta_0^c}\left(\boldsymbol{y}_k^{(j)}|\mathcal{H}_0\right)}\right]}_{\text{before change point}} \Rightarrow \underbrace{\ln\left[\frac{p_{\theta_1^c}\left(\boldsymbol{y}_k^{(j)}|\mathcal{H}_1\right)}{p_{\theta_1^c}\left(\boldsymbol{y}_k^{(j)}|\mathcal{H}_0\right)}\right]}_{\text{after change point}}$$

Simulation Setup (cont.)

We consider four clutter distributions:

- Gaussian distribution
- Student-t distribution
- K distribution
- Weibull distribution

Based on the proposed framework, 12 cases are included in the simulation results for target detection under nonstationary clutter as follows:

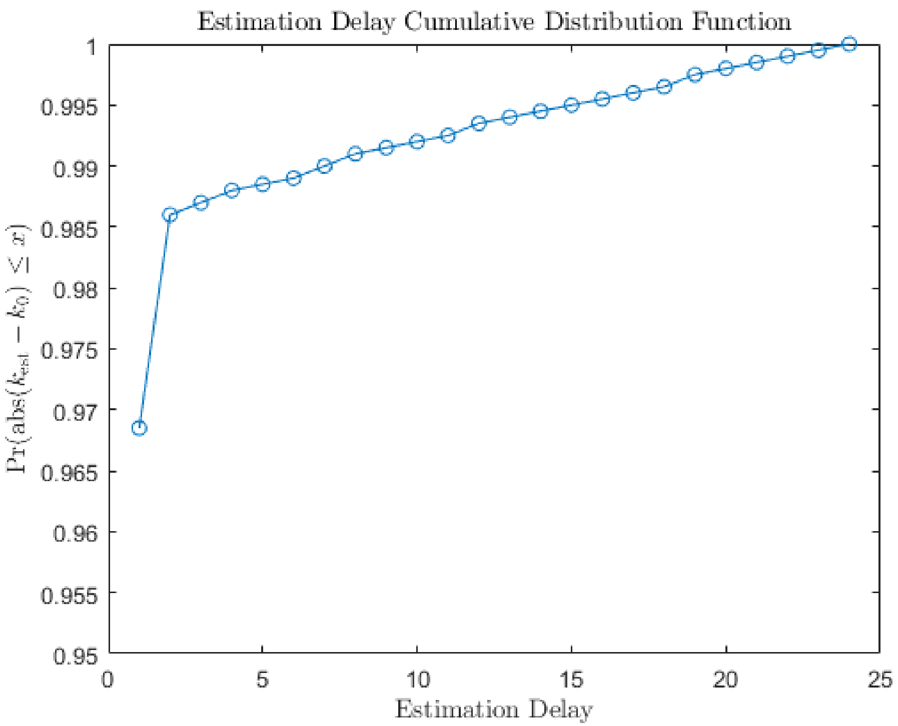
TABLE I: Distributions of Clutter in Target Detection

Before Change Point	After Change Point				
Gaussian		Student-t	K	Weibull	CUSUM
Student-t	Gaussian		K	Weibull	Extended CUSUM
K	Gaussian	Student-t		Weibull	Extended CUSUM
Weibull	Gaussian	Student-t	K		Extended CUSUM

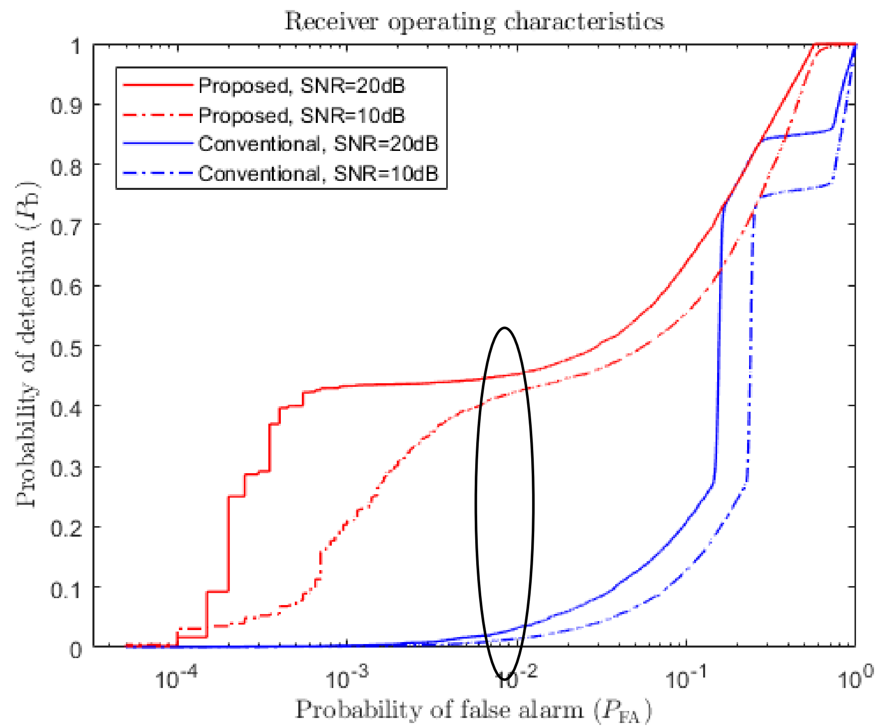
We present the results in terms of the (i) change point estimation delay and (ii) adaptive detection performance.

Numerical Results

Gaussian → Student-t



Change point detection (CUSUM)



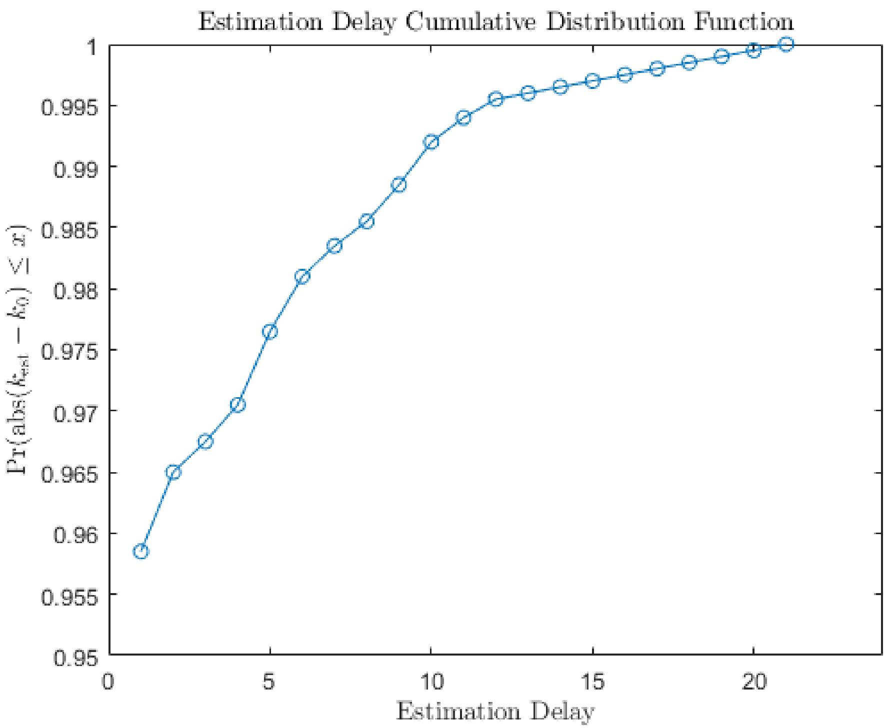
ROC curves (10dB and 20dB)

- Observation: Change in clutter distribution is detected within one processing interval for more than 96.5% of the time

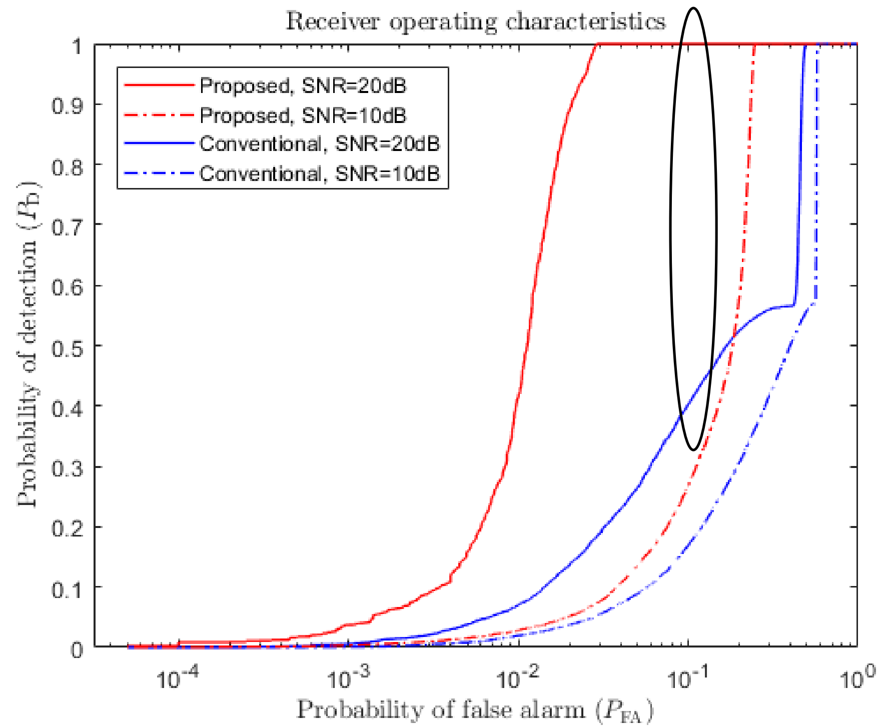
- Observation: Detection performance improves substantially (PD increased by ≈ 0.4) due to the incorporation of the modified clutter characteristics

Numerical Results (cont.)

Gaussian → K



Change point detection (CUSUM)

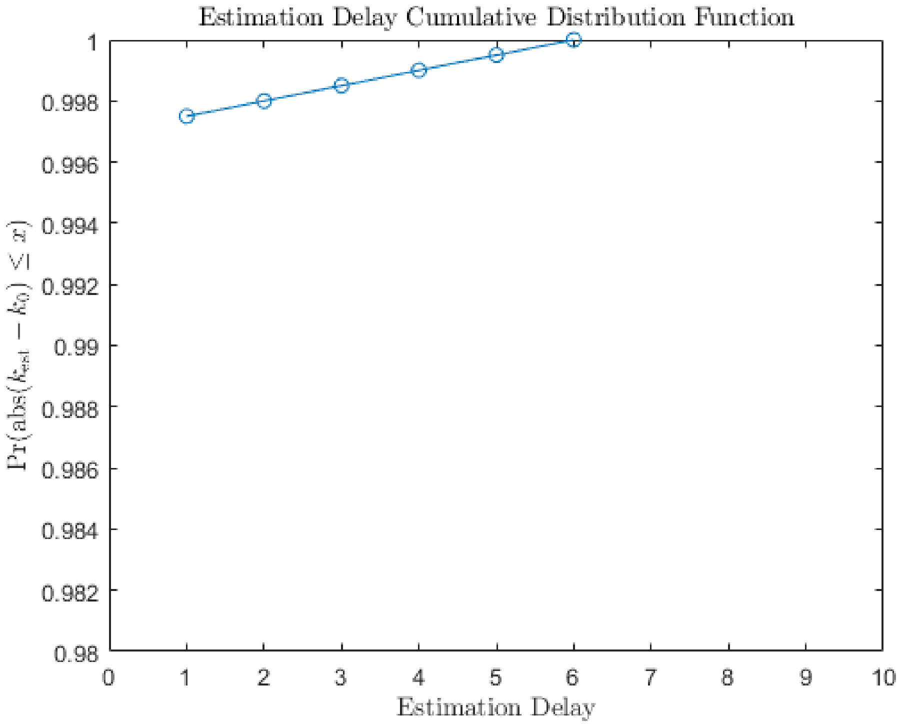


ROC curves (10dB and 20dB)

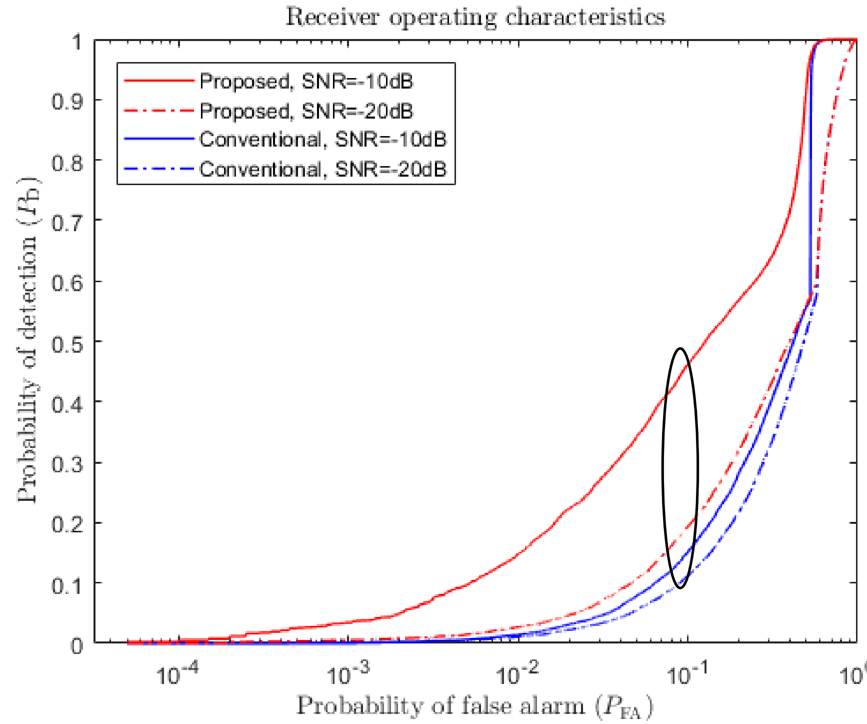
- Observation: Change in clutter distribution is detected within one processing interval for more than 95.5% of the time
- Observation: Detection performance improves substantially (PD increased by > 0.5) due to the incorporation of the modified clutter characteristics

Numerical Results (cont.)

Weibull → K



Change point detection (extended CUSUM)



ROC curves (-10dB and -20dB)

- Observation: Change in clutter distribution is detected within one processing interval for more than 99.6% of the time

- Observation: Detection performance improves substantially (PD increased by ≈ 0.4) due to the incorporation of the modified clutter characteristics

Numerical Results (cont.)

Change Point Detection Performance Summary:

Clutter before change point	Clutter after change point	Probability of change detection in one step delay	Estimation delay when probability of change detection ≥ 99%
Gaussian (G)	T	96.85%	7
	K	95.85%	10
	W	97.10%	6
Student-t (T)	G	91.80%	2
	K	99.40%	1
	W	98.90%	2
K	G	98.90%	3
	T	98.60%	2
	W	93.05%	2
Weibull (W)	G	97.20%	14
	T	97.00%	2
	K	99.75%	1

Numerical Results (cont.)

Improved Detection Performance Summary:

Clutter before change point	Clutter after change point	Improved P_D when $P_{FA} = 0.01$	
		SNR= 10dB	SNR= 20dB
Gaussian (G)	T	40.9%	42.1%
	K	0.8%	34.4%
	W	1.6%	48.7%
		SNR= -10dB	SNR= -20dB
Student-t (T)	G	4.0%	24.8%
	K	56.0%	11.7%
	W	64.4%	2.5%
K	G	*	*
	T	*	*
	W	86.3%	27.4%
Weibull (W)	G	*	*
	T	*	*
	K	13.4%	1.6%

* No comparable data exist since the case likelihood ratio=0/0 arise too often in conventional radar.

Conclusions and Future Work

Conclusions and Future Work

Conclusions:

- We developed a data-driven algorithm to detect a target in the presence of nonstationary environment (clutter).
- Applied an active drift learning technique to detect and estimate any change-point (if present) in the clutter distribution.
- Employed the incremental learning and drift detection algorithms to incrementally learn the environment and update the system parameters on the fly.
- Demonstrated with numerical examples that the proposed algorithm
  - quickly detects a change in the underlying clutter distribution,
  - significantly improves the target detection performance.

Future Work:

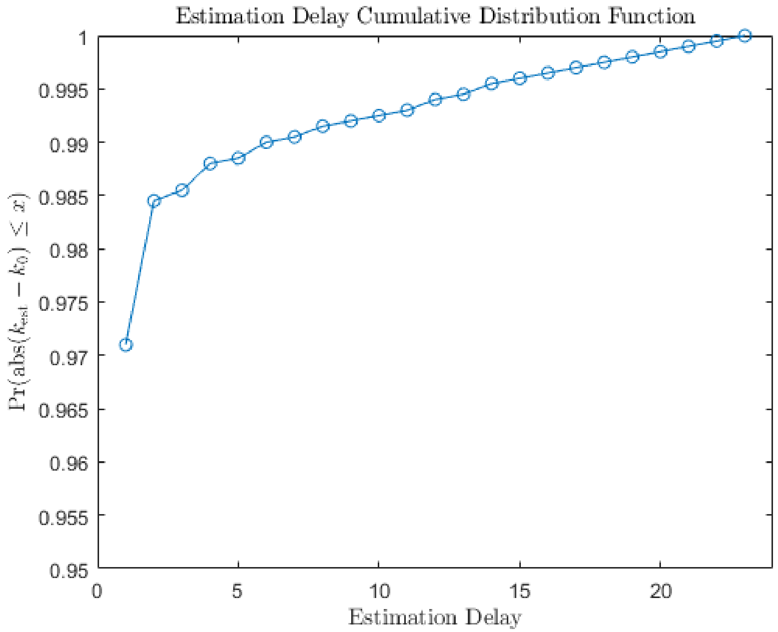
- Extend the model to include a Bayesian formulation of the change point parameter.
- Explore the active drift learning under noisy labels and passive drift leaning methodologies.
- Validate the performance of our proposed technique with real data.
- Incorporate waveform design and other adaptive techniques to further improve the detection and tracking performance under nonstationary environment.

Thank you!

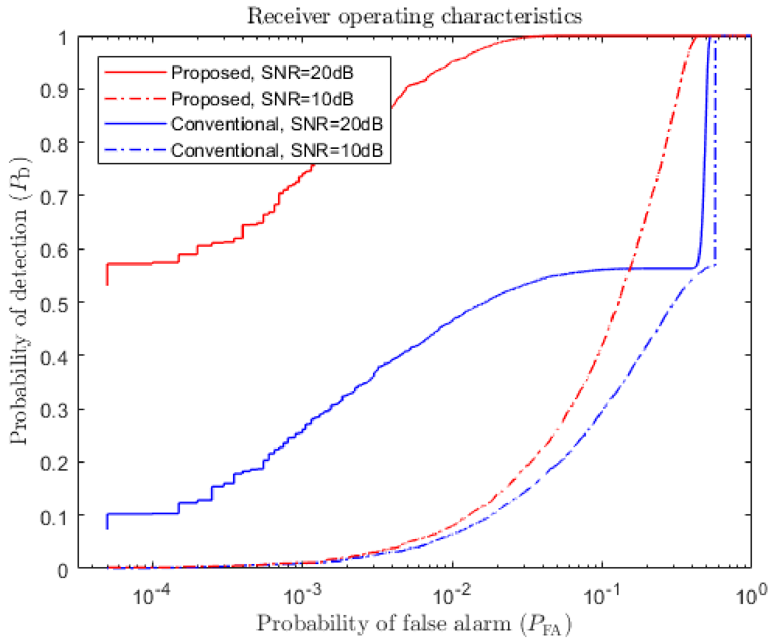
Appendix

More Numerical Results

Gaussian->Weibull



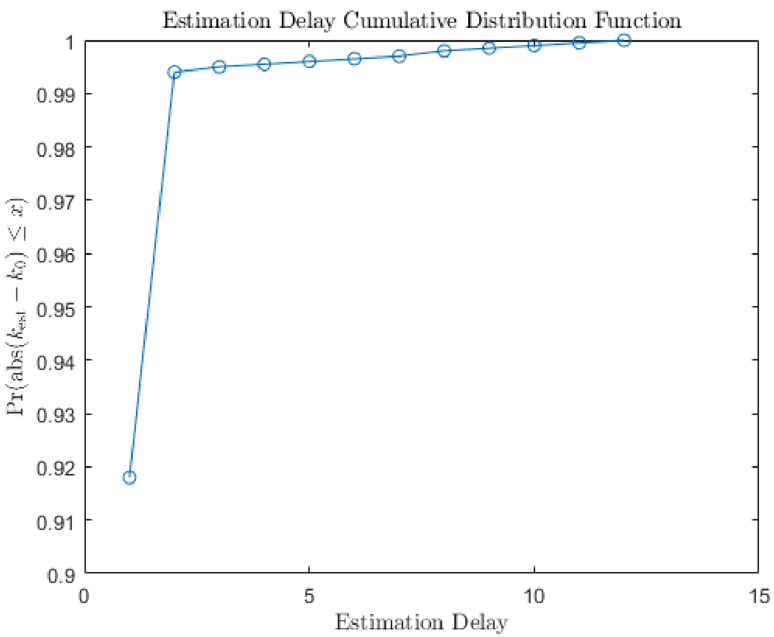
Change point detection (CUSUM)



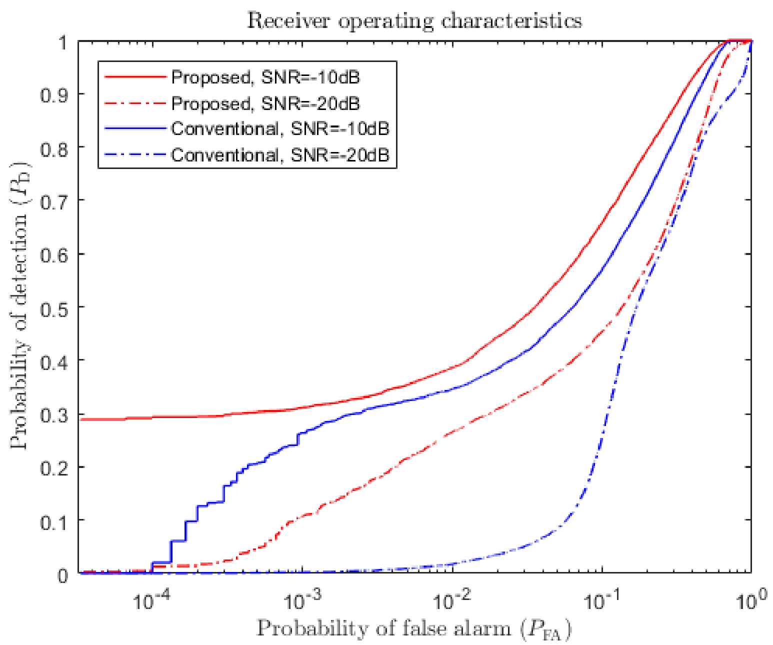
ROC curve (10dB and 20dB)

More Numerical Results

Student-t->Gaussian



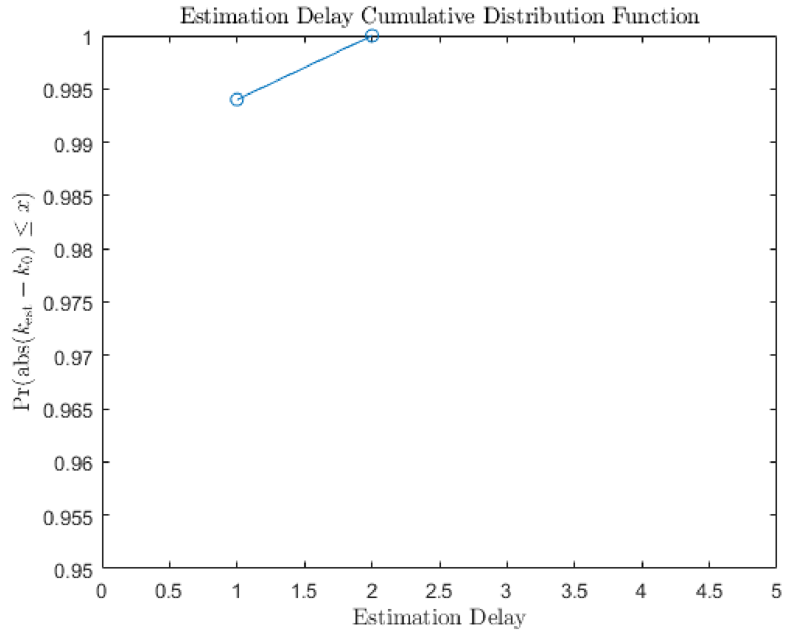
Change point detection (extended CUSUM)



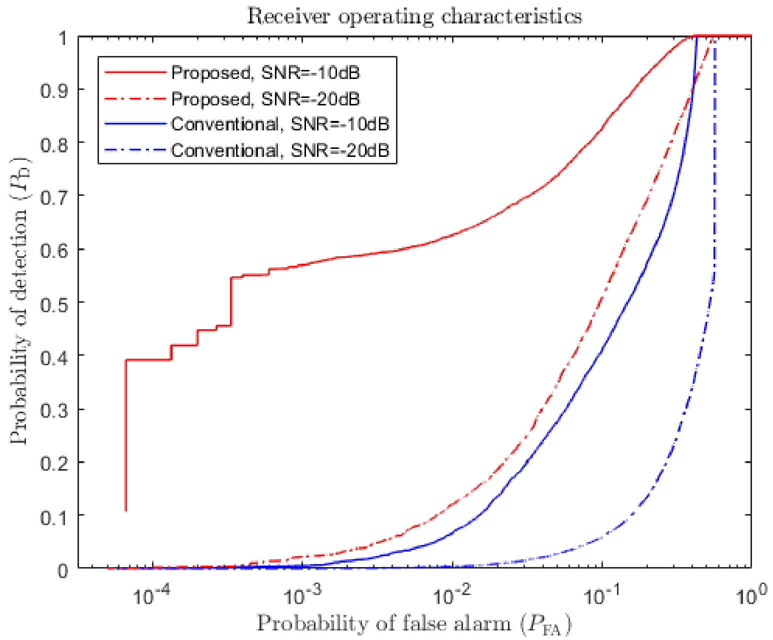
ROC curve (-10dB and -20dB)

More Numerical Results

Student-t->K



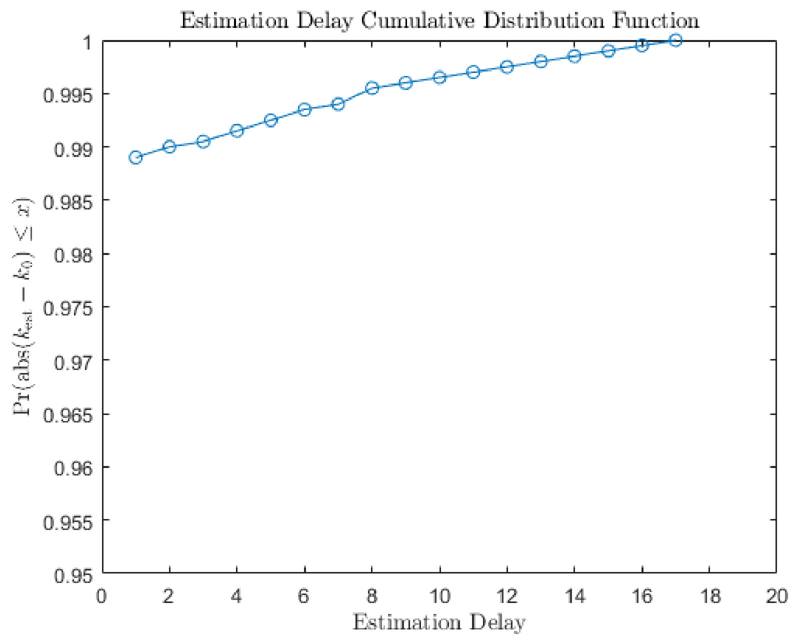
Change point detection (extended CUSUM)



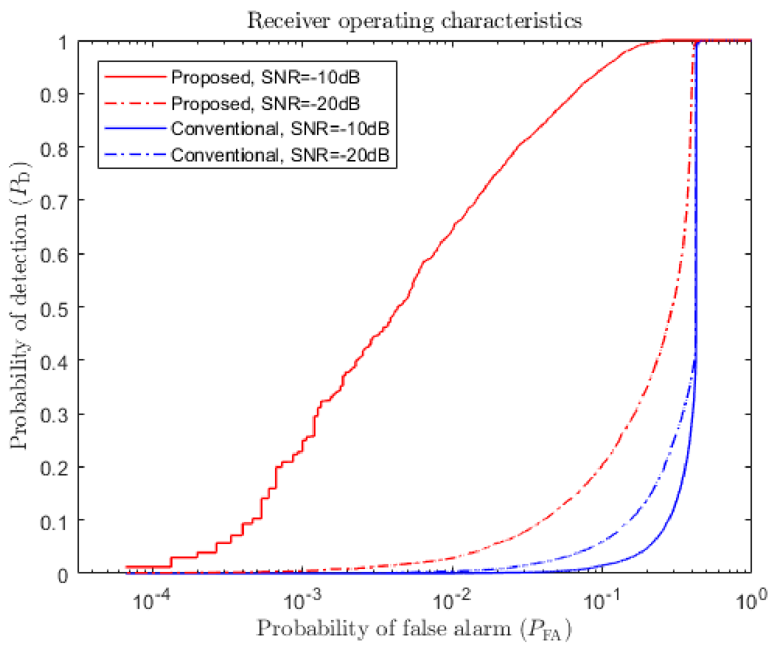
ROC curve (-10dB and -20dB)

More Numerical Results

Student-t->Weibull



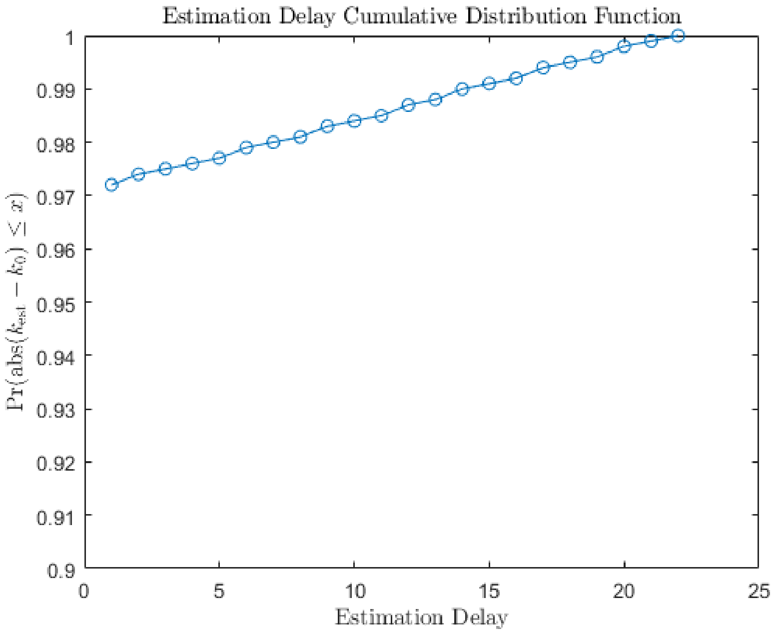
Change point detection (extended CUSUM)

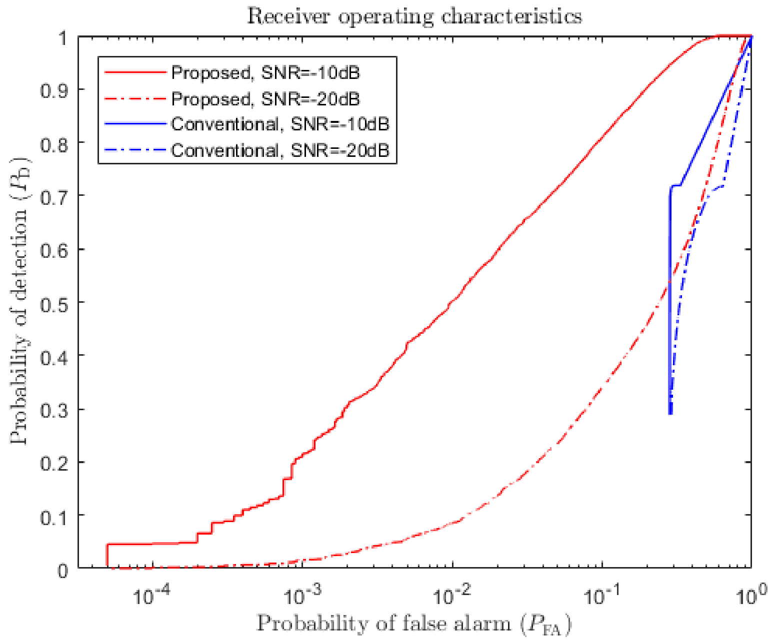


ROC curve (-10dB and -20dB)

More Numerical Results

Weibull->Gaussian



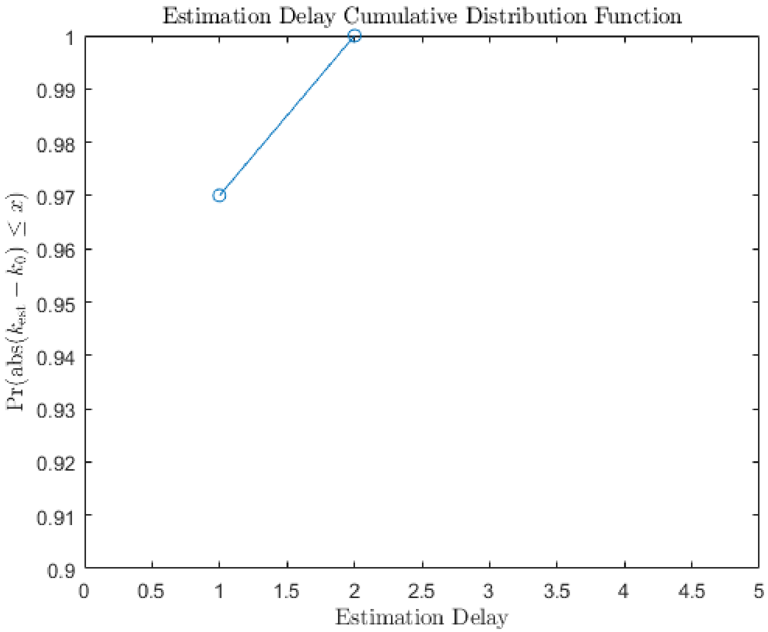


Change point detection (extended CUSUM)

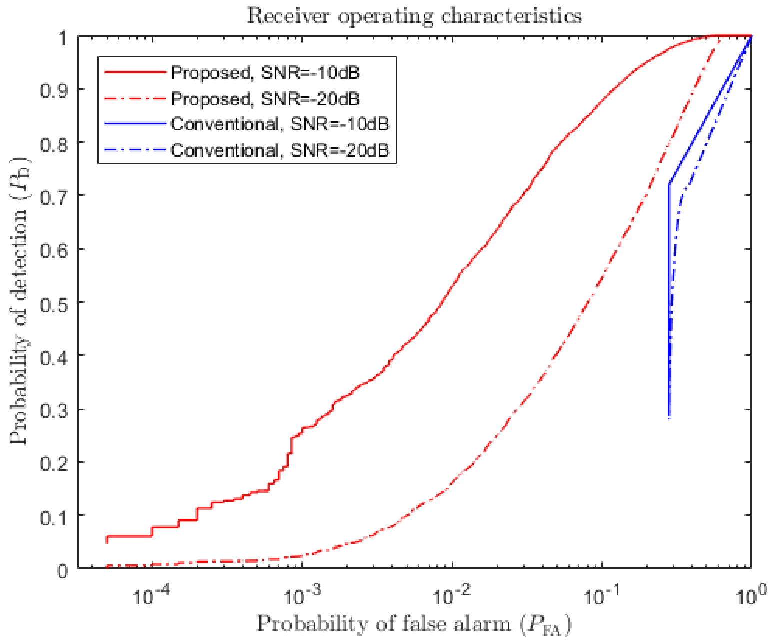
ROC curve (-10dB and -20dB)

More Numerical Results

Weibull->Student-t



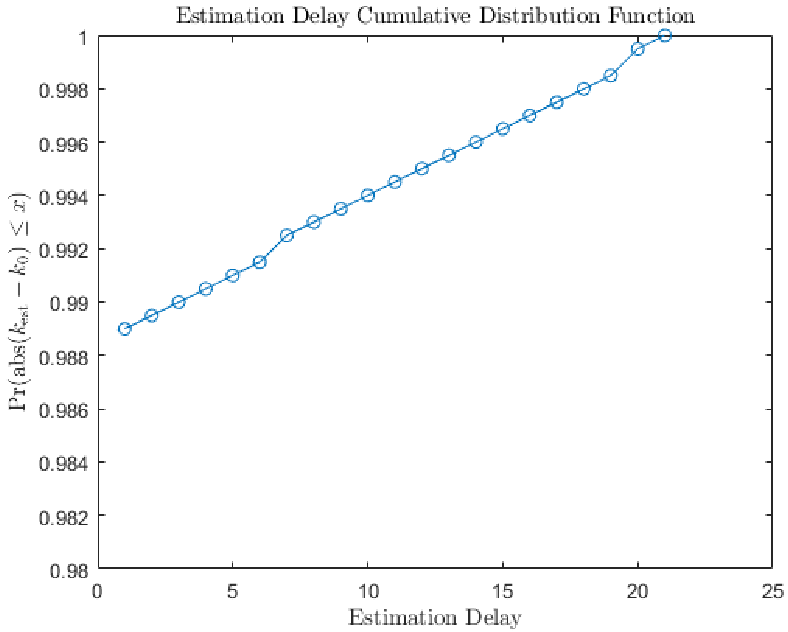
Change point detection (extended CUSUM)



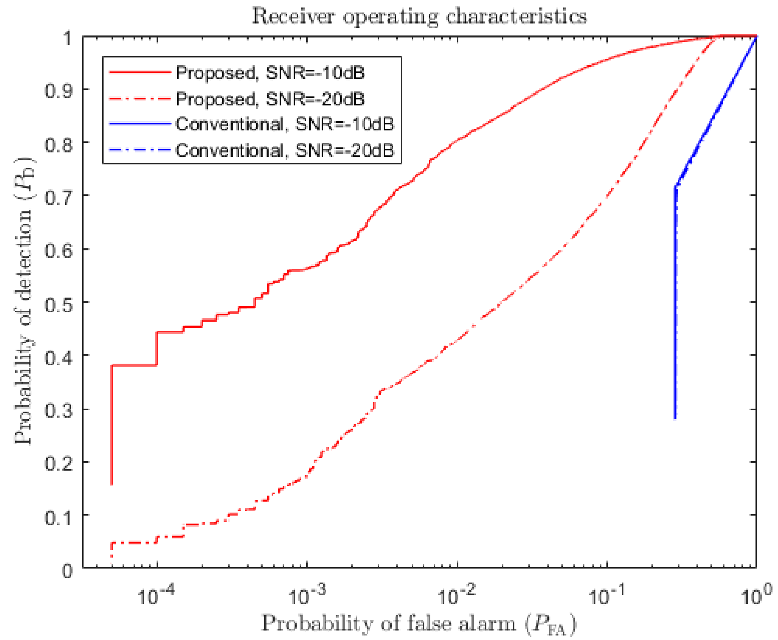
ROC curve (-10dB and -20dB)

More Numerical Results

K->Gaussian



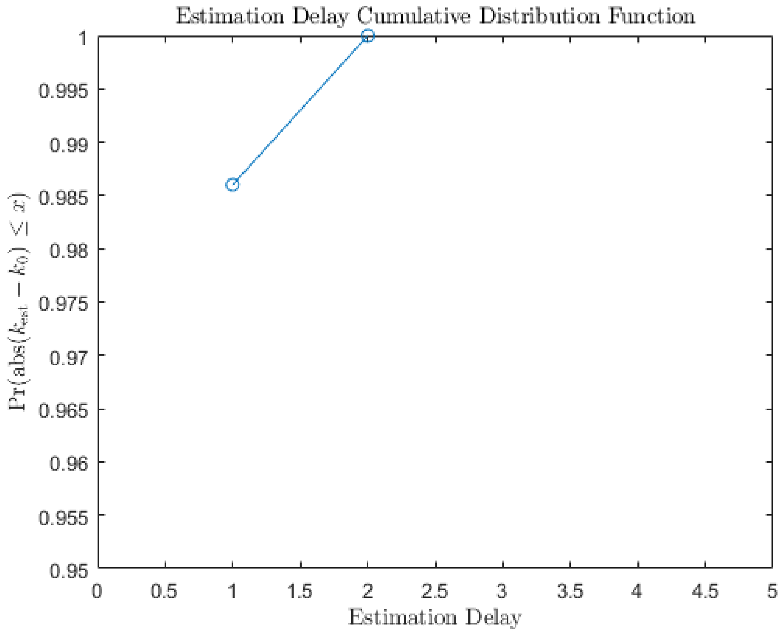
Change point detection (extended CUSUM)

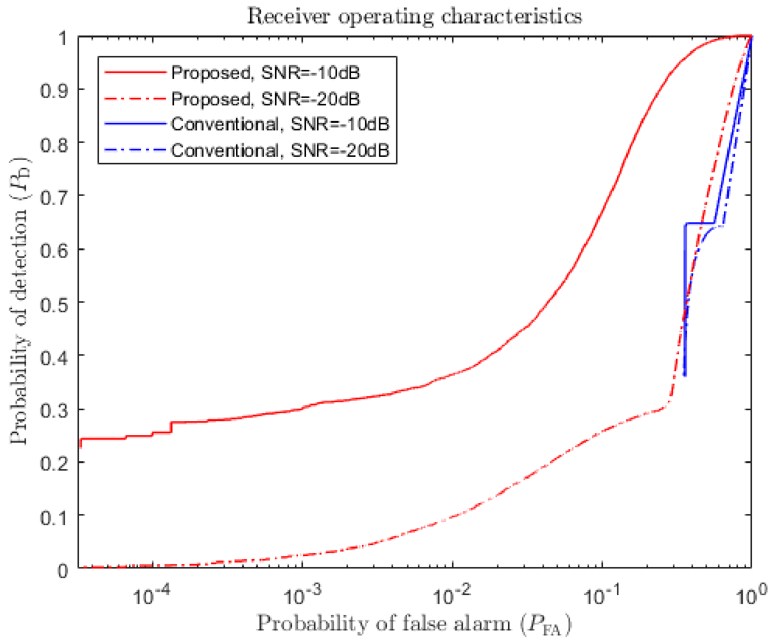


ROC curve (-10dB and -20dB)

More Numerical Results

K->Student-t



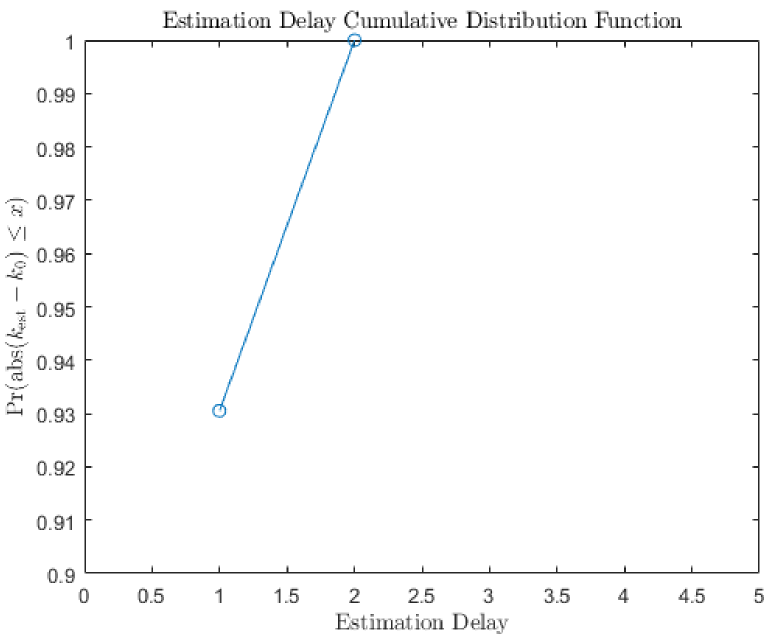


Change point detection (extended CUSUM)

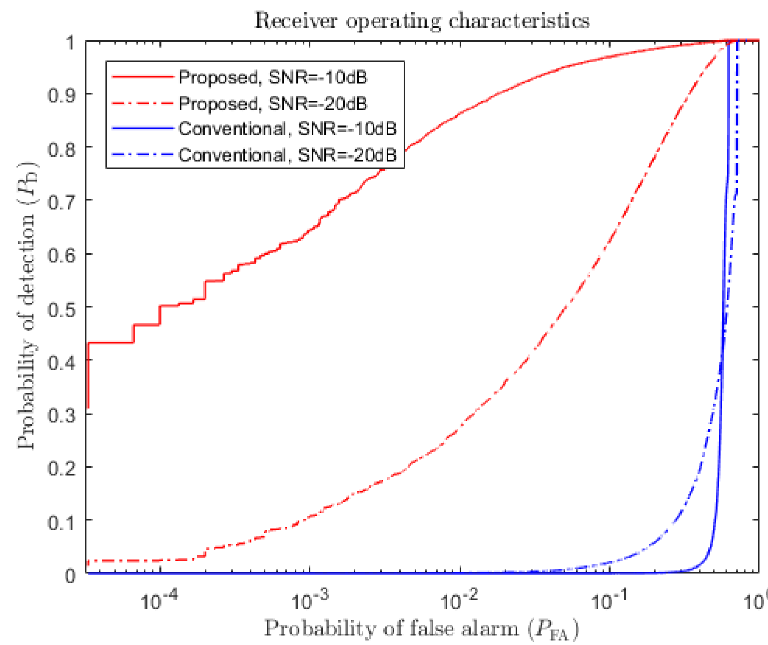
ROC curve (-10dB and -20dB)

More Numerical Results

K->Weibull



Change point detection (extended CUSUM)



ROC curve (-10dB and -20dB)

Numerical Results (cont.)

Improved Detection Performance Summary:

Clutter before change point	Clutter after change point	Improved P_D when $P_{FA} = 0.01$		Improved P_D when $P_{FA} = 0.001$	
		SNR=10dB	SNR=20dB	SNR=10dB	SNR=20dB
Gaussian (G)	T	40.9%	42.1%	20.8%	43.0%
	K	0.8%	34.4%	0.3%	3.2%
	W	1.6%	48.7%	0%	48%
		SNR=-10dB	SNR=-20dB	SNR=-10dB	SNR=-20dB
Student-t (T)	G	4.0%	24.8%	4.8%	10.6%
	K	56.0%	11.7%	56.5%	2.2%
	W	64.4%	2.5%	25.0%	0.5%
K	G	*	*	*	*
	T	*	*	*	*
	W	86.3%	27.4%	64.5%	10.1%
Weibull (W)	G	*	*	*	*
	T	*	*	*	*
	K	13.4%	1.6%	3.4%	0.4%

* No comparable data exist since the case likelihood ratio=0/0 arise too often in conventional radar.