Symposium SET-241 on 9th NATO Military Sensing Symposium

Adaptive Radar Signal Detection with Integrated Learning and Knowledge Exploitation

Prof. Hongbin Li
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ, USA

Dr. Muralidhar Rangaswamy
AFRL/RYAP
Bldg 620, 2241 Avionics Circle
WPAFB, OH 45433-7132, USA

May 31 to June 2, 2017

Supported by the AFOSR under Grant No. FA9550-16-1-0243

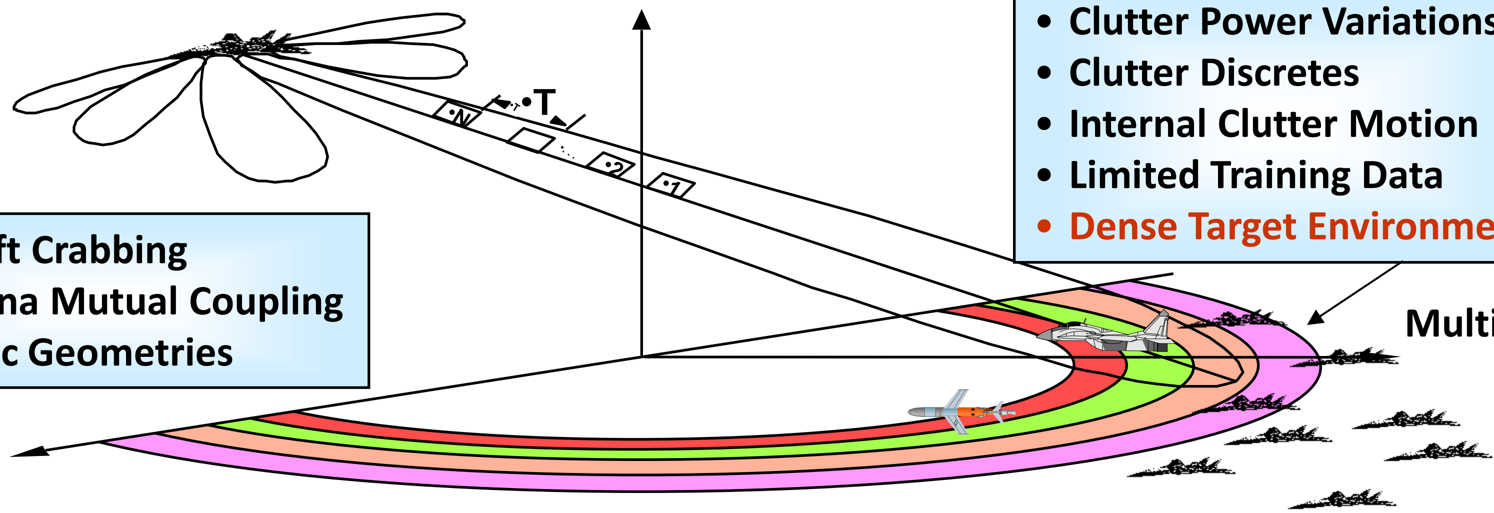
Outline

- Background the problem of interest
- Prior art on knowledge-aided target detection
- Proposed approach
  - Subspace clutter model and sparse representation
  - Bayesian framework
  - New hierarchical model for knowledge learning and exploitation
  - Bayesian inference
  - Numerical results
- Summary

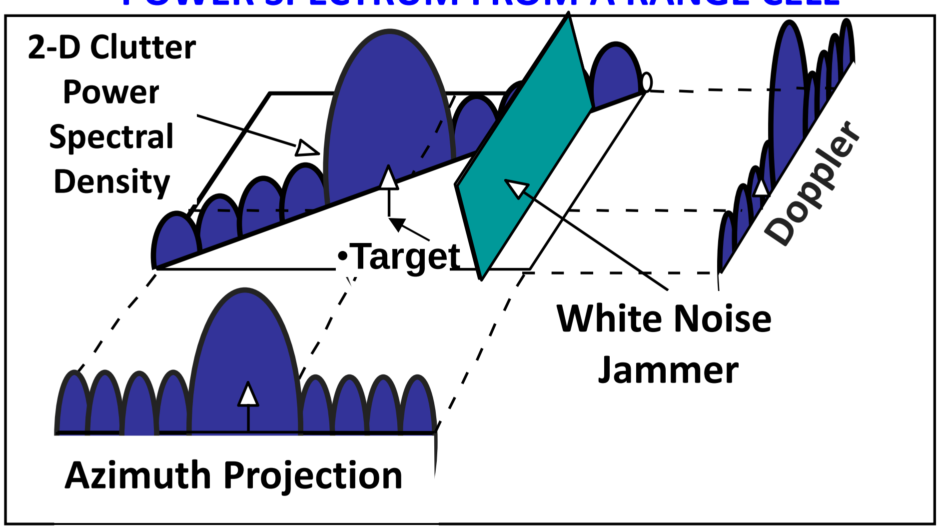
Airborne Radar Scenario

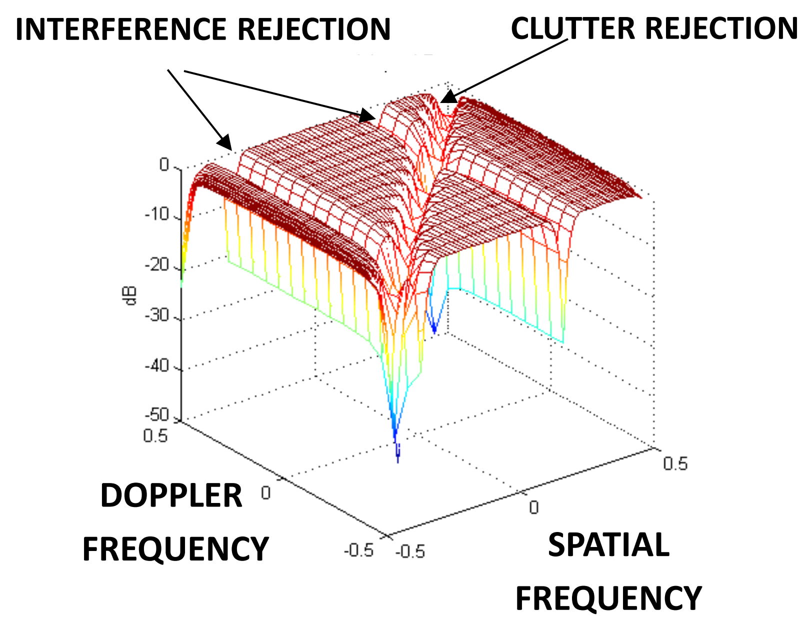
- Clutter Power Variations
- Clutter Discretes
- Internal Clutter Motion
- Limited Training Data
- Dense Target Environment

- Aircraft Crabbing
- Antenna Mutual Coupling
- Bistatic Geometries



POWER SPECTRUM FROM A RANGE CELL





M. Rangaswamy, SAM Challenges for Fully Adaptive Radar, 2012 IEEE SAM Workshop, Hoboken, NJ

The Problem

- An abstraction of the radar detection problem

$$H_0 : \mathbf{y} = \mathbf{d}$$
$$H_1 : \mathbf{y} = \alpha\mathbf{s} + \mathbf{d}$$

where

$\mathbf{y}$: $N \times 1$ test signal

α : unknown target amplitude

$\mathbf{s}$: $N \times 1$ target steering vector

$\mathbf{d}$: $N \times 1$ disturbance signal (clutter, noise, jamming, etc.)

- Above problem includes several important cases in radar detection
  - Beamforming: *N* = # of antennas
  - Doppler filtering: *N* = # of temporal pulses
  - Space-time adaptive processing (STAP): *N* = # of antennas x # of pulses
- The model can also be extended for MIMO radar signal detection

Knowledge-Aided Detectors

- Reduced-rank detectors still need considerable training data
- Parametric detectors, albeit requiring significantly less training data, may suffer from model mismatch
- **Knowledge-aided (KA) detectors:** Jointly exploit limited training/test data along with prior knowledge for detection (Guerci 2010)
  - Sources of prior knowledge: historical surveillance data, digital maps, physics/geometry based models of radar scene
  - Prior knowledge is often abstracted as a prior covariance matrix $\mathbf{R}_0$ of the disturbance/clutter signal $\mathbf{d}$
  - Key issue is to combine $\mathbf{R}_0$ with the data-dependent sample covariance matrix

$$\hat{\mathrm{R}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{y}_k \mathbf{y}_k^H$$

- Two approaches: deterministic vs. Bayesian combining

J. R. Guerci, *Cognitive Radar: The Knowledge-Aided Fully Adaptive Approach*, Artech House, 2010

Motivation

- In most existing KA approaches, prior knowledge takes the abstract form of covariance matrix. We are interested in exploiting knowledge in more natural format, such as
  - Spatial locations of dominant clutter scatterers (major natural and man-made objects in radar scene)
  - Angle-Doppler trace of clutter spectrum, which can be determined from mobility parameters of sensor platform (aircraft)
- Such info also represents knowledge at finer scales compared with the full-rank covariance matrix
- In addition, we hope to address the case when the prior knowledge is incomplete via Bayesian learning techniques
  - This takes into account clutter scatterers at locations that a priori unknown and helps improving the detection performance

Key Ideas of Proposed Approach

- Subspace model for disturbance signal
  - Location related knowledge naturally has a subspace structure
  - Radar clutter often has a low-rank subspace structure (Ward 1994)
- Sparse representation
  - Jamming/spatial interference usually arrives from a few angles
  - Clutter Doppler frequencies in ground based radar are close to zero with a small spread
  - In airborne STAP systems, clutter exhibits a clutter ridge on the angle-Doppler domain
- New hierarchical models for knowledge exploitation
- Above components are integrated in a Bayesian learning framework for radar detection

J. Ward, "Space-time adaptive processing for airborne radar," Technical Report 1015, Lincoln Laboratory, MIT, Dec 1994

Subspace Model and Sparse Representation

- Recall the subspace model for the disturbance (clutter, jamming, noise)

$$\mathbf{y} = \mathbf{H}\boldsymbol{\beta} + \mathbf{n}, \quad \mathbf{H} \in \mathbb{C}^{N \times L} (L < N), \quad \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

  - Both subspace basis matrix **H** and coefficient vector **β** are unknown
  - Simultaneous estimation of both **H** and **β** is a nonlinear problem
- The problem can be linearized by using an over-determined dictionary matrix **A** and a sparse constraint:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$$

$\mathbf{A} : N \times M$ dictionary matrix $(N \ll M)$

$\mathbf{x} : M \times 1$ sparse coefficient

- A number of approaches (greedy methods, convex optimization, log-sum minimization, etc.) can be used to solve the above sparse recovery problem. We will employ a Bayesian framework to facilitate learning and knowledge exploitation

Bayesian Learning Framework

- A popular Bayesian approach to sparse recovery is based on the Gaussian-inverse Gamma hierarchical prior. Sparse vector **x** has a Gaussian prior:

$$p(\mathbf{x}|\boldsymbol{\alpha}) = \prod_{m=1}^{M} p(x_m|\alpha_m), \quad p(x_m|\alpha_m) = \mathcal{N}(x_m|0, \alpha_m^{-1})$$

- Inverse variance parameter α has a Gamma distribution:

$$p(\boldsymbol{\alpha}) = \prod_{m=1}^{M} \text{Gamma}(\alpha_m|a,b) = \prod_{m=1}^{M} \Gamma(a)^{-1} b^a \alpha_m^{a-1} e^{-b\alpha_m}$$

  a and *b* are hyperparamters set to very small values, e.g., 10^{-4}, to provide broad non-informative hyperprior for α

- A broad hyperprior allows the posterior mean of α_m to become unbounded, driving x_m to 0 and leads to a sparse estimate of **x** (Tipping 2001)
- Above Bayesian formulation does not employ any prior knowledge on the support of **x** (equivalently, the subspace basis **H**)

M.E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *The Journal of Machine Learning Research*, pp. 211-244, 2001

2-Layer Hierarchical Model for Knowledge Integration

- If the hyperparamter *b* takes a relatively large value, the inverse Gamma hyperprior no longer encourages large values of α_m. Therefore, we should use different b_m for different α_m, depending on prior knowledge
- Suppose we have partial knowledge of the support of the sparse vector **x**

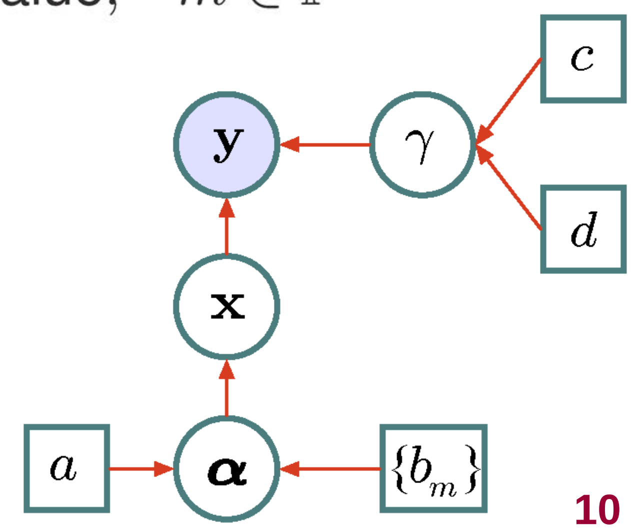
$$\mathbb{P} \subset \{m | x_m \neq 0, m = 1, 2, \ldots\}$$

which denotes a subset containing indices of some non-zero x_m

- We propose the following hyperprior and hyperparameters

$$p(\boldsymbol{\alpha}) = \prod_{m=1}^{M} \text{Gamma}(\alpha_m | a, b_m), \quad b_m = \begin{cases} \text{large value}, & m \in \mathbb{P} \\ \text{small value}, & m \in \mathbb{P}^c \end{cases}$$

- Large b_m over the known support helps to promote non-sparse solution over the support (and benefit from prior knowledge)
- Small b_m over the unknown support region helps recovering missing subspace bases and maintain sparsity



Bayesian Inference

- The problem is to estimate parameter associated with the proposed model and recover the sparse signal **x** from **y** using the posterior

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})}, \qquad \boldsymbol{\theta} \triangleq \{\mathbf{x}, \boldsymbol{\alpha}, \gamma\} \quad \text{for 2-layer model}$$

- The posterior cannot be computed analytically due to the difficulty in computing the marginal distribution $p(\mathbf{y})$
- We use variational Bayesian methods based on mean field approximation to compute the posterior

$$p(\boldsymbol{\theta}|\mathbf{y}) \approx \prod_{k=1}^{K} q_k(\boldsymbol{\theta}_k),$$

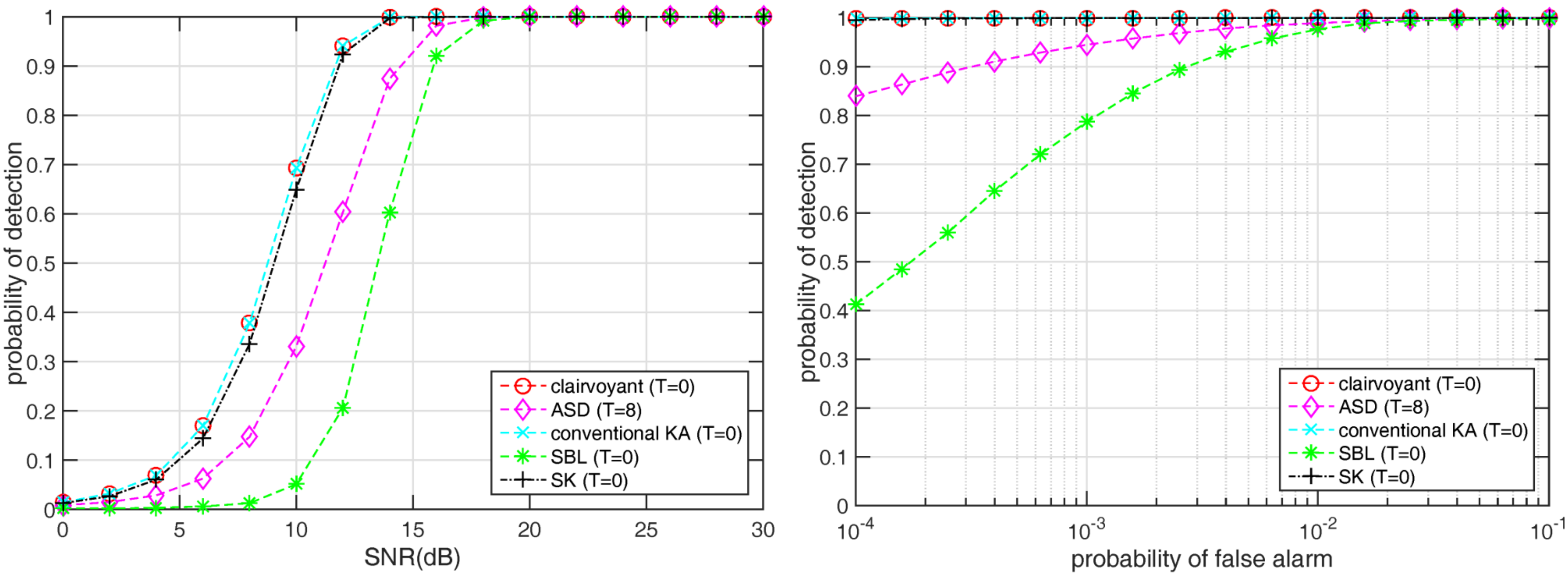
$$q_k(\boldsymbol{\theta}_k) = \frac{\exp\left(\left\langle \ln p(\mathbf{y}, \boldsymbol{\theta}) \right\rangle_{l \neq k}\right)}{\int \exp\left(\left\langle \ln p(\mathbf{y}, \boldsymbol{\theta}) \right\rangle_{l \neq k}\right) d\boldsymbol{\theta}_k}$$

- The resulting detector is referred to as subspace knowledge (SK) aided detector

Numerical Results

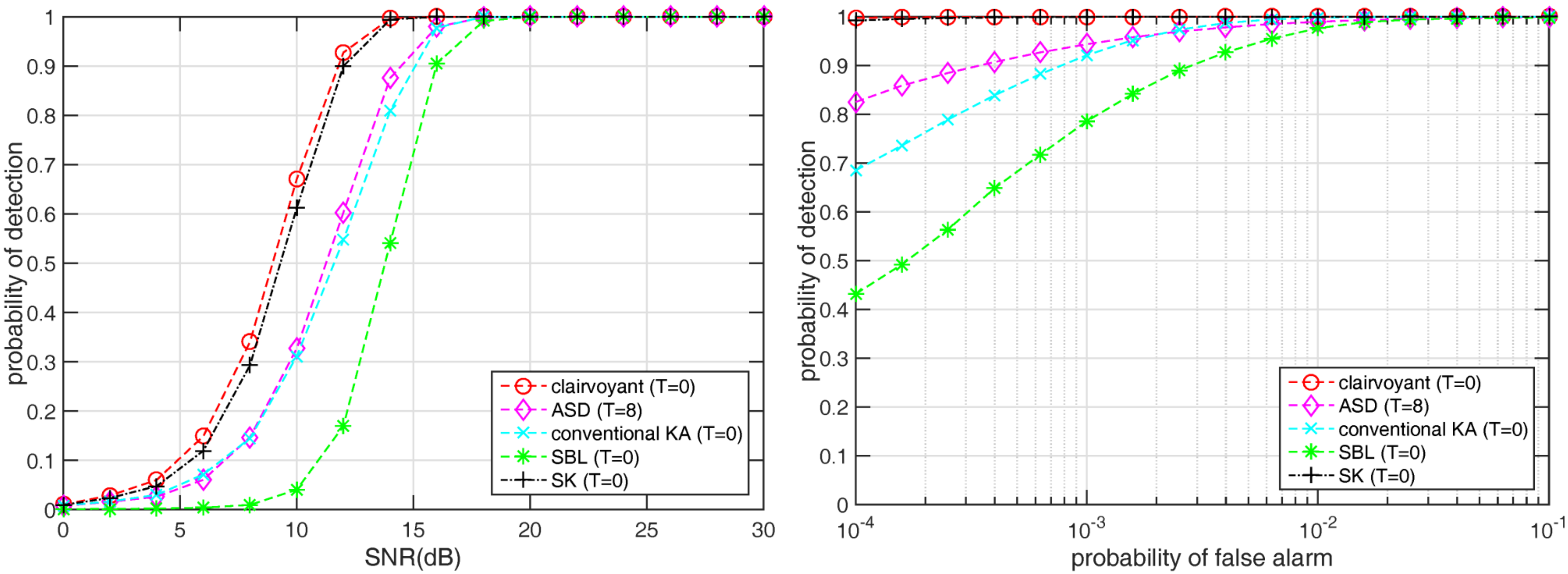
- Two cases are considered for the proposed SK detector
  - Case 1: disturbance subspace is fully known (ideal scenario)
  - Case 2: disturbance subspace is partially known (only 4 out of the 7 Fourier vectors) are assumed known
- We compare our SK detector with
  - Clairvoyant subspace detector with full knowledge
  - Conventional KA (knowledge-aided) detector that relies on the prior knowledge for disturbance mitigation
  - ASD (adaptive subspace detector) that uses training data to estimate the disturbance subspace
  - SBL (sparse Bayesian learning) based detector
- Only ASD requires training data ($T > 0$) while the other detector do not ($T = 0$)

Case I (Full Knowledge)



Left: probability of detection P_d vs. SNR with INR = 30 dB and $P_f = 10^{-3}$.
Right: ROC curve with SNR = 15 dB and INR = 30 dB

Case 2 (Partial Knowledge)



Left: probability of detection P_d vs. SNR with INR = 30 dB and $P_f = 10^{-3}$.
Right: ROC curve with SNR = 15 dB and INR = 30 dB

Summary

- This work is concerned with the fundamental problem of detecting weak signals in strong interference
  - Non-homogeneous environments
  - Limited training data and prior knowledge
- Proposed approach builds on recent advances in compressed sensing, sparse signal processing, machine learning
- Specific achievements
  - New sparsity and hierarchical Bayesian models which can handle incomplete prior knowledge
  - New hypothesis testing solutions for GMTI and other similar applications

STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY

Thank you!