AN RFI DETECTION ALGORITHM FOR MICROWAVE RADIOMETERS USING SPARSE COMPONENT ANALYSIS

Priscilla N. Mohammed[1,2], Asmita Korde-Patel[1], Armen Gholian[1], Jeffrey R. Piepmeier[1], Adam Schoenwald[1], Damon Bradley[1]

[1]NASA's Goddard Space Flight Center, Greenbelt, MD, USA
[2]Morgan State University, Baltimore, MD, USA

Introduction

Radio frequency interference (RFI) is a problem for microwave remote sensing of Earth. Although frequency allocations are set aside for passive sensing, RFI can still degrade measurement quality. In some cases radiometer bandwidth exceeds allocated spectrum to reduce measurement uncertainty or spectrum allocations are shared, forcing microwave radiometer to co-exist with terrestrial sources. Low level RFI is particularly detrimental as it can be concealed as natural variability leading to flawed scientific results. RFI detection algorithms have been developed to address the problem. Research into other algorithms is needed to improve upon the sensitivity of existing detection algorithms to various types of RFI. The Sparse Component Analysis (SCA) has been investigated to determine its sensitivity to continuous wave (CW) RFI.

Sparse Component Analysis (SCA)

SCA is a blind source separation method which seeks to extract *N* unknown sources from *P* observations where *P* < *N*. The sources need to have disjoint supports.

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t), \qquad t = 1, \dots, T$$

observations (known) — mixing matrix (unknown) — sources (unknown)

$$\mathcal{H}_0: \begin{cases} x_H = a_{11}n_H + a_{12}n_V \\ x_V = a_{21}n_H + a_{22}n_V \end{cases}$$

No RFI

$$\begin{pmatrix} x_H \\ x_V \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} n_H \\ n_V \\ r \end{pmatrix}$$

$$\mathcal{H}_1: \begin{cases} x_H = a_{11}n_H + a_{12}n_V + a_{13}r \\ x_V = a_{21}n_H + a_{22}n_V + a_{23}r \end{cases}$$

With RFI

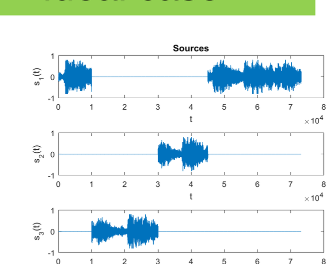
n_H Radiometer horizontal polarization

n_V Radiometer vertical polarization

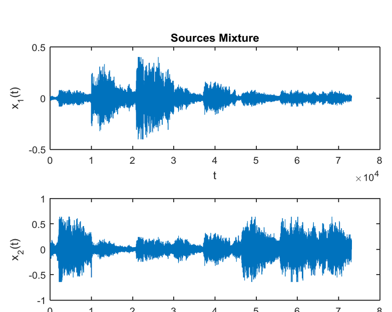
r Radio frequency interference (RFI)

Methodology

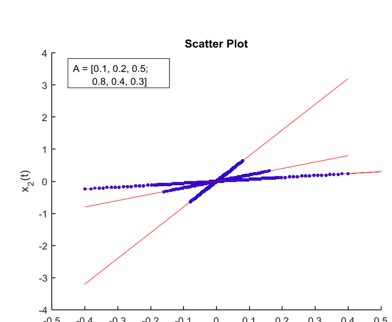
Ideal case



Sources with disjoint supports



Sources mixture

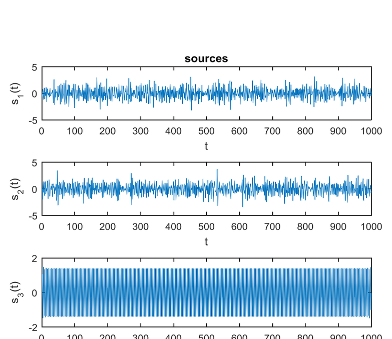


Scatter plot of the mixtures x_1 and x_2

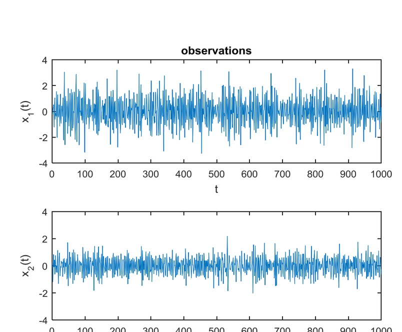
Use least squares to estimate columns of mixing matrix, A

$$\text{Least squares: } \hat{s}_n(t) = \begin{cases} \dfrac{<x(t), \hat{A}_n>}{\|\hat{A}_n\|^2} & t \in \hat{\Lambda}_n \\ 0 & otherwise \end{cases}$$

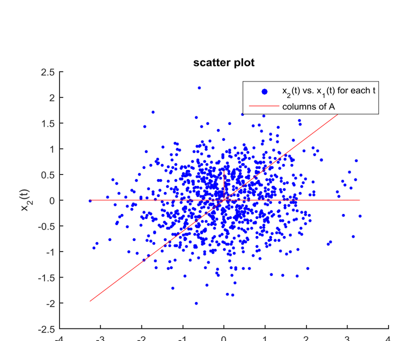
In practice the source signals are not disjoint in time. Columns of A not easily determined from scatter plot



Sources with non disjoint supports

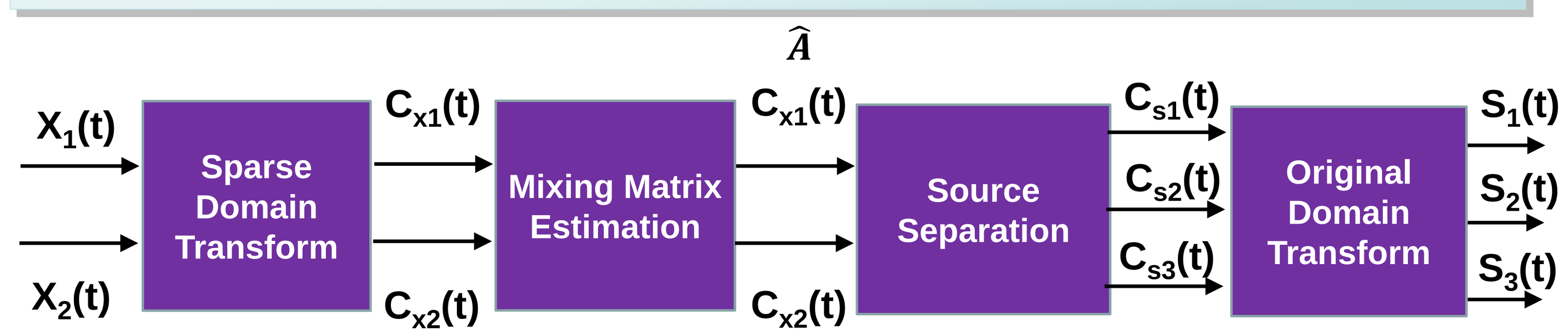


Sources mixture



Scatter plot of the mixtures x_1 and x_2

SCA in Practice



- Each source is transformed to a sparse representation in another domain.
- Mixing matrix $\boldsymbol{A}$ is estimated using the sparse coefficients obtained in the transformation.
- The sources are separated in the transformed domain using the coefficients of observations and the estimated mixing matrix $\boldsymbol{A}$.
- The sources are finally reconstructed in the time domain.

Algorithms
- Dictionary
  - Dictionary learning
  - Structured dictionary
- Joint sparse representation
  - Global optimization
  - Greedy algorithm
- Mixing matrix estimation
  - Global clustering
  - Local scatter plots
- Separation
  - Binary masking

RFI Detection

- The detection criterion is the median of the absolute value of the reconstructed sources, $\hat{x}$ in time.
- The output of SCA are three reconstructed sources in time, $\hat{s_1}(t), \hat{s_2}(t), \hat{s_3}(t)$ where t = 1, 2, 3 … *N*.
- The median of the absolute value of each reconstructed source ($median(|\hat{s_n}(t)|), n = 1, 2, 3$) is evaluated. If all medians are greater than a given threshold, RFI is present.

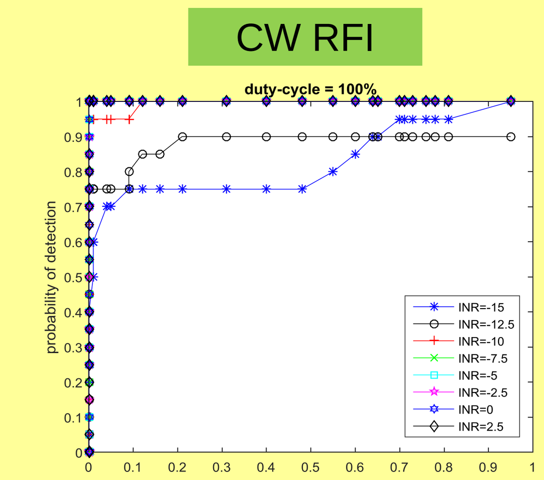
Simulation Model

$$\begin{pmatrix} x_H \\ x_V \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} n_H \\ n_V \\ r \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \sqrt{INR_H} \\ 0 & a_{22} & \sqrt{INR_V}a_{22} \end{pmatrix}$$

and $INR_V = INR_H \pm 3dB$

CW RFI



Sparse Domain Transform

Represent each source by a linear combination of a few elementary signals (atoms): $s(t) = \sum_{k=1}^{K} c_s(k)\varphi_k(t)$

Measure of sparsity

$$\|\boldsymbol{c}_s\|_\tau = \begin{cases} \left(\sum_{k=1}^{K} |c_s(k)|^\tau\right)^{1/\tau} & \tau > 0 \\ \text{num of nonzero coefficients} & \tau = 0 \end{cases}$$ quantifies the sparsity of $\boldsymbol{c}_s$

Ideal τ for sparsity is 0

Dictionary

Dictionary: set of elementary signals known as atoms $\boldsymbol{\Phi}$

Consider dictionaries that span C^T, $K > T$: redundant dictionary, infinite representation

Once the dictionary is determined the signal representation can be determined in the transformed domain using either a global optimization technique or greedy algorithms.

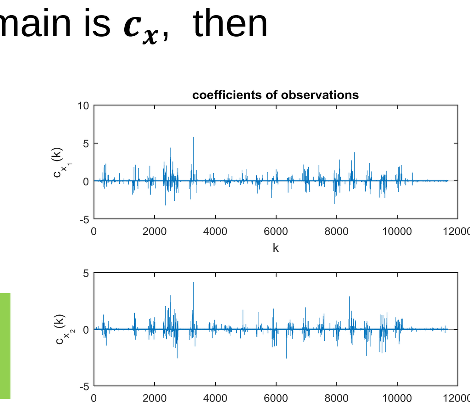
Let K, denote the number of atoms in the dictionary, $\boldsymbol{\Phi}$. If the signal in time domain is $\boldsymbol{x}$ and the signal coefficients in the transformed domain is $\boldsymbol{c}_x$, then

$$\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{c}_x$$

where $\boldsymbol{x} = [x(1), \dots, x(t)]^T$ and $\boldsymbol{c}_x = [c_x(1), \dots, c_x(K)]^T$.

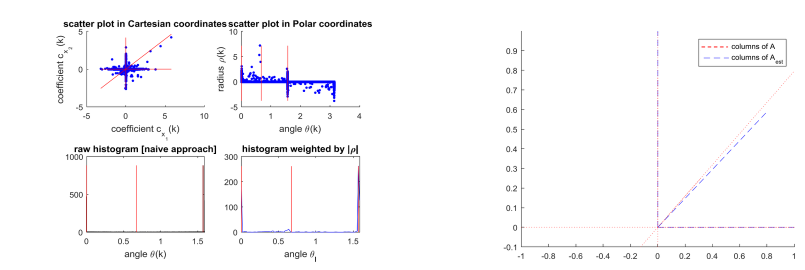
$$\boldsymbol{c}_x = \arg \min_{\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{c}_x} \|\boldsymbol{c}_x\|_1$$

Sparse Representation of sources mixture



Estimating Mixing Matrix

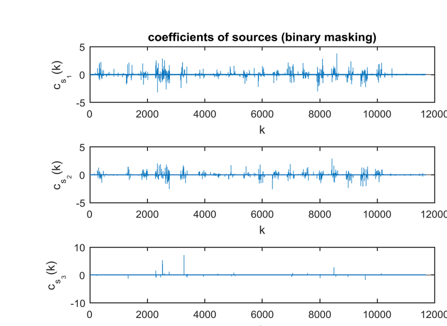
Scatter plot of coefficients $\boldsymbol{c}_x$



- The mixing matrix $\boldsymbol{A}$ is estimated through the scatter plot of the coefficients $\boldsymbol{c}_x$.
- A global clustering algorithm, weighted histogram, is used to estimate the directions of the columns of the mixing matrix.

Separation and Reconstruction

- Binary masking is used to separate the sources.
- A mask is derived by setting the indices of the sources which correspond to the highest amplitude in observations given the mixing matrix with a value of one.
- Since the signals are disjoint, only the active sources are assigned a coefficient value, while the other sources are masked with zeroes.
- In the last step, the acquired source signals can be reconstructed back to their original domain by using $\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{c}_x$.



Separation
Coefficients of sources

Source 1

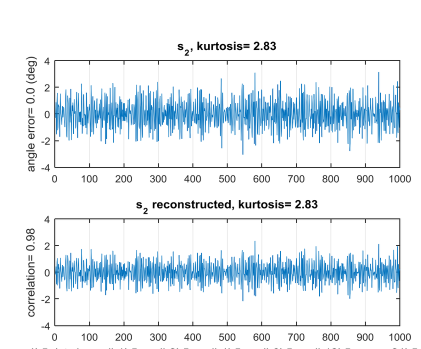
Source 2

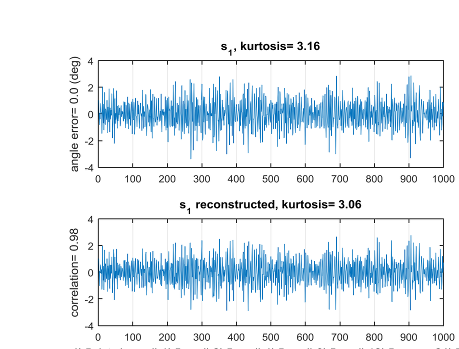
S1 actual

S2 actual

S1 reconstructed

S2 reconstructed





Conclusion

- Monte Carlo simulations with 1000 time samples are used, along with the structured dictionary and Orthogonal matching pursuit (OMP) algorithm for joint sparse representation, the weighted histogram for matrix estimation and binary masking for source separation.
- Figure shows the performance results of SCA for detection of CW RFI with INR_H ranging from -15 dB to 2.5 dB.
- The results show perfect to near perfect detection for INRs greater than -12.5 dB and very good detection at -12.5 dB and -15 dB.
- Results show that detection works for relatively large INRs for CW RFI.

Reference

R. Gribonval and M. Zibulevsky. "Sparse component analysis," in *Handbook of Blind Source Separation: Independent component analysis and applications* P. Comon, C. Jutter, Academic press, 2010, pp. 367-420.