Nonstationary Signal Analysis and Array Processing

Yimin D. Zhang, Ph.D.

Wireless Communications & Positioning Lab
Center for Advanced Communications
Villanova University
Nonstationary Frequency Modulated (FM) Signals

Examples of instantaneous narrowband nonstationary signals:

- Linear FM (LFM, chirp)
- Polynomial phase signal
- Nonparametric FM signals

They are frequently encountered in many radar and other applications:

- Radar & sonar sources (LFM)
- Target return (Doppler)
- Smart jammer
- Others (birds, whales...)

FM-CW radar

OTH radar return

GPS jammer (anti-tracker)
Motion Classification in Through-the-Wall Radar

Two arm swing  One arm swing  No arm swing

(c) Y. D. Zhang
Old-age dependency ratio (population aged 65-years or over to that aged 20-64) to reach 30% in 2020

Direct medical costs for fall-related senior injuries reach $55 billion by 2020

Radar-based technology is attractive
- Non-invasive monitoring
- Privacy
-Insensitive to light and object obstruction
- Low-cost

Fall detection heavily relies on time-frequency analysis
Doppler Signatures in Assisted Living

Stand up and sit down

Walking without cane

Walking with cane

Fall
GPS Smart Jamming

- Nonstationary jammers are difficult to remove by a single domain mitigation algorithm.
- Joint time-frequency domain algorithm removes jammers with minimum distortion to GPS signal.
- Jammer estimation and excision becomes much more challenging there existing missing data samples.
Various applications of nonstationary signals demand:

- Instantaneous frequency (IF) estimation
- Waveform reconstruction and removal
- Detection and classification
- DOA estimation
- Separation of signals with close IF signatures
- All these for signals with undersampled data

This presentation focuses on nonparametric FM signals characterized by their IF signatures.
Outline of presentation

1: Time-Frequency Analysis of FM Signals and Applications in DOA Estimation and Tracking
   - Time-Frequency Representations
   - Spatial Time-Frequency Distribution
   - Time-Frequency MUSIC for DOA Estimation
   - Separation of Closely Spaced FM Signals

2: Sparse Reconstruction of FM Signals from Observations with Missing Samples
   - Effect of Missing Data in TFD
   - Adaptive Kernel Design and Sparse Reconstruction
   - Multi-Sensor TFD Reconstruction and DOA Estimation
Outline of presentation

1: Time-Frequency Analysis of FM Signals and Applications in DOA Estimation and Tracking

- Time-Frequency Representations
- Spatial Time-Frequency Distribution
- Time-Frequency MUSIC for DOA Estimation
- Separation of Closely Spaced FM Signals

2: Sparse Reconstruction of FM Signals from Observations with Missing Samples

- Effect of Missing Data in TFD
- Adaptive Kernel Design and Sparse Reconstruction
- Multi-Sensor TFD Reconstruction and DOA Estimation
Time-Frequency Analysis

Classical time-frequency analysis techniques can be classified as

- Linear - e.g., short-time Fourier transform (STFT), fractional Fourier transform
- Bilinear - e.g., Wigner-Ville distribution, Cohen’s class

STFT: Simple but generally does not provide high resolution

\[
\text{STFT}(n, k) = \sum_{m=-N/2}^{N/2-1} x(n + m)w(m)e^{-j \frac{2\pi}{N} mk}
\]
Instantaneous autocorrelation function (IAF)

\[ R_{xx}(t, \tau) = x(t + \tau)x^*(t - \tau) \]

Wigner-Ville distribution (WVD)

\[ D_{xx}(t, f) = \sum_{\tau = -\infty}^{\infty} x(t + \tau)x^*(t - \tau)e^{-j4\pi f \tau} \]

\[ = \text{DFT}_\tau[R_{xx}(t, \tau)] \]
Bilinear Time-Frequency Distribution

- Bilinear time-frequency distribution introduces the undesired cross-terms.

- A number of approaches has been developed to reduce the cross-terms.

- Cohen’s class of reduced-interference distribution (RID): applies a kernel in the ambiguity domain.

- Ambiguity function

\[ A_{xx}(\theta, \tau) = \sum_{t=-\infty}^{\infty} x(t + \tau)x^*(t - \tau)e^{-j4\pi \theta t} = \text{DFT}_t[R_{xx}(t, \tau)] \]
Signal Representation in Different Domains

Waveform

Time-frequency distribution

Ambiguity function

FFT (f, τ)

FFT (θ, t)

Cross-terms

Auto-terms

Doppler θ

Delay τ

Frequency f

Time t

Waveform - real part

FFT (f, τ)

FFT (θ, t)
Time-Frequency Kernels

- Auto-terms are located around the origin of the ambiguity domain, and cross-terms tend to be away from the origin.
- Time-frequency kernel is a low-pass filter in the ambiguity domain.
- The kernel $\phi(m, \tau)$ defines different types of TFDs:
  - Data-independent: e.g., Choi-Williams distribution
  - Data-dependent: e.g., adaptive optimal kernel
Time-Frequency Analysis

Wigner-Ville distribution

Choi-William distribution

STFT

Adaptive optimal-kernel

(c) Y. D. Zhang
Array Signal Processing

Far-field narrowband FM signal $d(t)$

\[ \theta \]

$N$ sensors

\[ l \]

\[ \text{Steering vector} \]

\[ \phi = 2\pi \left( \frac{l}{\lambda} \right) \sin(\theta) \]

\[ x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{bmatrix} = d(t) \begin{bmatrix} 1 \\ \exp(j\phi) \\ \vdots \\ \exp(j(N-1)\phi) \end{bmatrix} = d(t) \mathbf{a}(\theta) \]
Array Signal Model

General form with $P$ signals and additive noise vector

$$x(t) = y(t) + n(t) = A(\Theta)d(t) + n(t)$$

where

- $x(t)$ : array signal vector $(N \times 1)$
- $y(t)$ : noise-free array signal vector $(N \times 1)$
- $d(t)$ : signal vector $(P \times 1)$
- $A(\Theta)$ : steering matrix $(N \times P)$
- $n(t)$ : additive noise vector $(N \times 1)$
STFD vs. Covariance Matrix

Covariance matrix of $x(t)$

$$R_{xx} = E[x(t)x^H(t)] = AR_{dd}A^H + R_{nn}$$

Definition of spatial time-frequency distribution (STFD) matrix:

$$D_{xx}(t, f) = \text{DFT}_{2D}\{\text{DFT}_t[x(t + \tau)x^H(t - \tau)]\}$$

where the cross-TFD between $x_i(t)$ and $x_j(t)$ is

$$[D_{xx}(t, f)]_{i,k} = D_{x_ix_k}(t, f) = \text{DFT}_{2D}\{\text{DFT}_t[x_i(t + \tau)x_k^*(t - \tau)]\}$$

for $i, j=1, \ldots, P$. 
STFD vs. Covariance Matrix

Covariance matrix of \( x(t) \)

\[
R_{xx} = E[x(t)x^H(t)] = AR_{dd}A^H + R_{nn}
\]

STFD matrix of \( x(t) \) at a \((t, f)\) point

\[
D_{xx}(t, f) = AD_{dd}(t, f)A^H + D_{nn}(t, f)
\]

STFD matrix of \( x(t) \) averaged over a T-F region \( \Omega \)

\[
D_{xx}(\Omega) = A \left[ \sum_{(t, f) \in \Omega} D_{dd}(t, f) \right] A^H + \sum_{(t, f) \in \Omega} D_{nn}(t, f)
\]

Same relationship:

Covariance matrix \( \leftrightarrow A \rightarrow \) source covariance matrix

STFD matrix \( \leftrightarrow A \rightarrow \) source TFD matrix

**Advantage:** STFD is defined in the selected T-F region, allowing signal **selection** and **enhancement.**
Consider selecting $P_0$ FM signals out of $P$.

STFD matrix averaged over the autoterm points defined over $P_0$ signals

$$D_{xx} = A^o \left[ \frac{(L/n_o)}{R_{dd}} \right] (A^o)^H + \sigma I$$

$\sigma$ : noise power
$L$ : length of window used in pseudo WVD
$A^o, R^o_{dd}$ : defined at $P_o$ selected signals

Key advantages of STFD:

- SNR enhancement
- Signal discrimination
- Process more sources than array sensors
Time-Frequency MUSIC

Conventional MUSIC

MUSIC algorithm estimates the DOAs by finding the $P$ largest peaks of the localization function

$$f(\theta) = \left| \hat{\mathbf{G}}^H \mathbf{a}(\theta) \right|^{-2}$$

$\hat{\mathbf{G}}$: noise subspace of covariance matrix

Time-frequency MUSIC

Time-frequency MUSIC algorithm estimates the DOAs by finding the $P_o$ largest peaks of the localization function

$$f(\theta) = \left| \left( \hat{\mathbf{G}}^{tf} \right)^H \mathbf{a}(\theta) \right|^{-2}$$

$\hat{\mathbf{G}}^{tf}$: noise subspace of STFD matrix
Effect of Signal Enhancement

- MUSIC performance comparison (8 sensors, SNR=-20 dB, \( L = 129 \))
- Advantages of enhanced SNR are limited to low SNR because of sample coherence.
Effect of Signal Discrimination

- With source discriminations in T-F domain, we can
  - achieve significant performance improvement for closely spaced sources
  - handle more sources than sensors

MUSIC performance comparison
(8 sensors, SNR= -5 dB, L=129)
Over-the-Horizon Radar challenges

- **Dynamic frequency range: 3 – 30 MHz**
  - Challenges in antenna and array design
  - Mutual coupling & impedance matching
  - Degrees-of-freedom of waveforms

- **Narrowband Signals: typical BW 10 – 50 KHz**
  - Low resolution (typical cell size: 10 km x 10 km)
  - Doppler signature analysis

- **Low SCR and SNR**
  - Mitigation of ground/sea clutter

- **Multipath Propagation**
  - Suppression of unstable F-layer multipath
  - Target association
  - A major challenge is the estimation of target altitudes.
  - One way is through the Doppler analysis of the local multipath.
When $R >> H >> h$, \[ l_1 \approx R + \frac{2H^2 - 2Hh}{R}, \quad l_2 \approx R + \frac{2H^2 + 2Hh}{R} \]
Example: a maneuvering aircraft makes a 180° turn in 30.72 sec to change height and direction.

The time interval corresponds to 6 revisits, and each revisit contains 256 samples.

The frequency difference between the three signatures reveals the target speed in the elevation direction.
The relative altitude estimation amounts to Doppler difference estimation. The challenges lie in

- Nonlinear Doppler
- Close Doppler separation
Knowledge-based processing:
Turn challenges into opportunities!

- Nonlinear Doppler
  > to be stationarized

- Close Doppler separation
  > can be stationarized together

- Known number of components and their symmetry
Signal Estimation Based on Warping

- The signal is segmented into multiple half-overlapping short-time periods.
- In each segment, the Doppler is characterized using a local polynomial phase modeling of order 3.
- Parameters are estimated using multilag high-order ambiguity function (mlHAF).
Target Trajectory & Doppler Signature

Top view

Side view

(c) Y. D. Zhang
Received signal vector at the array

\[ y(t) = \sum_{i=1}^{L} \sigma_s \rho_i e^{j \psi_i(t)} a_i + z_i \]

where

- \( L \): Number of multipaths
- \( \sigma_s \): Transmit signal magnitude
- \( \rho_i \): Radar cross section (RCS) for the \( i \)th path
- \( \psi_i \): Phase term of the \( i \)th path (mainly due to target Doppler)
- \( a_i \): steering vector for the \( i \)th path
- \( z_i \): additive Gaussian noise
Signal Filtering MUSIC

Stationarization with the estimated phase of the kth signal

\[ y(t)e^{-j\hat{\psi}_k(t)} = \sum_{i=1}^{L} \sigma_s \rho_i e^{j\psi_i(t)-j\hat{\psi}_k(t)} a_i + z_i e^{-j\hat{\psi}_k(t)} \]

\[ \approx \sigma_s \rho_k a_k + \sum_{\substack{i=1 \\text{to} \ L \ \text{L}}_{i \neq k}} \sigma_s \rho_i e^{j\psi_i(t)-j\hat{\psi}_k(t)} a_i + z_i e^{-j\hat{\psi}_k(t)} \]

Weighted summation (i.e., LPF to only keep the selected component) for DOA estimation:

\[ \overline{y}^{[k]} \approx \sigma_s \rho_k a_k + \overline{z}_i^{[k]} \]

Example: assuming idea IF estimation

Only estimate DOD/DOA for this signal
Altitude Estimation through Direction Finding

MIMO-based

6 Tx antennas, 10 Rx antennas
(minimum redundancy array)
input SNR = -10 dB
Proposed Technique

Extended Kalman filter (EKF) is tailored to solve this problem:

- Obtained individual array data for each component
- Resolved Doppler signatures are used as part of the measurement data
- The maximum a-posteriori (MAP) criterion is used to estimate the initial altitude and motion direction.

Hypothesis about target altitude and direction
Resolved Doppler signature estimates
Resolved hypothesis uncertainty

Extended Kalman Filter (EKF)

Output estimate of state

Prior knowledge of state

Measurements $y_k$
Simulation Results (SNR=-10 dB)

(a) Target ground range

(b) Target altitude

(c) Range-direction velocity

(d) Elevation velocity
Conclusion – Part I

- The spatial time-frequency distribution (STFD) provides a natural framework for improved array processing of nonstationary signals.
- The time-frequency signature characterization can be used for source discrimination and SNR enhancement.
- TF-MUSIC was shown as an example for DOA estimation with source discrimination capability.
- Signal stationarization is useful to resolve closely separated Doppler signatures and used for DOA estimation and target tracking.
Outline of presentation

1: Time-Frequency Analysis of FM Signals and Applications in DOA Estimation and Tracking
   - Time-Frequency Representations
   - Spatial Time-Frequency Distribution
   - Time-Frequency MUSIC for DOA Estimation
   - Separation of Closely Spaced FM Signals

2: Sparse Reconstruction of FM Signals from Observations with Missing Samples
   - Effect of Missing Data in TFD
   - Adaptive Kernel Design and Sparse Reconstruction
   - Multi-Sensor TFD Reconstruction and DOA Estimation
Introduction

- Missing data may arise from
  - Line-of-sight obstruction and/or propagation fading
  - Removal of data samples contaminated by impulsive noise
  - Intentional undersampling for reduced hardware complexity

- Missing data samples yield missing entries in the IAF and produces noise-like artifacts in TFD.

- These artifacts can be mitigated by applying time-frequency kernels or using data interpolation.

- Sparse signal reconstruction methods improve performance over the direct application of Fourier transform.
Signal Model

- Consider $K$ narrowband FM signals impinging on an array consisting of $N$ sensors
  \[ y(t) = A(\Theta)d(t) + n(t), t = 1, \ldots, T \]

- Consider a thinned sampling of the array observations with a random pattern applied to each array sensor
  \[ x_p(t) = y_p(t) \cdot b_p(t) \]

  where

  \[ b_p(t) = \begin{cases} 
  1, & \text{if } t \in S_p \\
  0, & \text{if } t \notin S_p 
  \end{cases} \]

  \[ S_p \subset \{1, \ldots, T\} : \text{set of observed time instants with cardinality } |S_p| = T - M_p \]

  \[ M_p : \text{number of missing samples} \]
Effect of Missing Data

Waveform  TFD (WVD)  Ambiguity function  IAF

Full data (noise-free)

50% of missing data (noise-free)
Missing data samples yield spreading artifacts that are randomly distributed over the entire T-F domain, and the overall variance increases as the number of missing data samples increases.

For T-F points where $W_{xx}(t,f)$ is zero or insignificant, the variance is uniformly distributed over $f$, whereas the variance depends on $t$ because of the zero-padding effect.
Outline of presentation

- The artifacts due to missing data samples
  - Spread over the entire T-F and AF domain
  - Resembles that due to noise
  - Can be mitigated through a proper T-F kernel

- T-F Kernels
  - Mitigate both artifacts and cross-terms
  - Best kernels keep the signal signature and filter out other regions
  - Signal-adaptive kernels (e.g., AOK) are desirable
**Data-Dependent Kernel**

- **Adaptive Optimal Kernel (AOK)**
  - Kernel is important in time-frequency analysis to suppress cross-terms while preserving auto-terms
  - AOK is a well-known **data-dependent** kernel, which is obtained by solving the following optimization problem defined in the polar coordinate system:

    \[
    \max_{\psi} \int_{0}^{2\pi} \int_{0}^{\infty} |A(r, \psi)\Phi(r, \psi)|^2 r dr d\psi \\
    \text{subject to } \Phi(r, \psi) = \exp\left(-\frac{r^2}{2\sigma(\psi)}\right) \\
    \frac{1}{4\pi^2} \int_{0}^{2\pi} \int_{0}^{\infty} |\Phi(r, \psi)|^2 r dr d\psi = \frac{1}{4\pi^2} \int_{0}^{2\pi} \sigma(\psi) d\psi \leq \alpha
    \]

Data-Dependent Kernel

- AOK substantially mitigates the missing data artifacts and clearly shows auto-term characteristics.
- Localized Choi-Williams distribution (CWD) kernels emphasize the locality and yield missing or weak TFD entries around missing data positions.
Sparsity-Aware AOK

- Notice the vertical strips (impulsive noise) due to missing data.

- Sparsity-aware AOK:

\[
\begin{align*}
\text{minimize} & \quad L_2(\sigma) + \lambda L_1(\sigma) \\
\text{subject to} & \quad \Phi(r, \psi) = e^{-\frac{r^2}{2\sigma^2(\psi)}}, \\
& \quad \frac{1}{2\pi} \int \sigma^2(\psi) d\psi \leq \alpha,
\end{align*}
\]

where

\[
L_2(\sigma) = \iint r dr d\psi |A(r, \psi) - A(r, \psi) \Phi(r, \psi)|^2
\]

and

\[
L_1(\sigma) = \iiint r \rho \sin(\psi - \phi) d\rho d\psi d\phi
\]

Sparse Reconstruction

- TFDs can be reconstructed based on their sparsity in the T-F domain.

- AF and TFD are related by 2-D DFT matrix:
  - Large-dimensional dictionary matrix
  - Global sparse reconstruction (cannot specify $t$ to perform local reconstruction)

Sparse Reconstruction

• We use the 1-D DFT relationship between IAF and TFD:

\[ c^{[t]} = \Phi w^{[t]} + \epsilon^{[t]} \]

• TFD reconstruction can be performed locally for each \( t \).
Structure-Aware Bayesian Compressive Sensing

- Bayesian CS maximizes the posterior probability over all unknown parameters.
- To encourage the TF signature sparsity, we place a spike-and-slab prior to $\mathbf{w}^{[t]}$:

$$p(\mathbf{w}_l^{[t]}|\pi_l^{[t]}, \beta_0^{[t]}) = (1 - \pi_l^{[t]})\delta(\mathbf{w}_l^{[t]}) + \pi_l^{[t]}\mathcal{N}(\mathbf{w}_l^{[t]}|0, [\beta_0^{[t]}]^{-1})$$

for frequency domain index $l \in [1, ..., T]$, where

- $\pi_l^{[t]}$: prior probability of a nonzero element
- $\beta_0^{[t]}$: precision of Gaussian distribution

- Let $\mathbf{w}_l^{[t]} = z_j^{[t]} \theta_l^{[t]}$ to make the inference analytical, where $\theta_l^{[t]} \sim \mathcal{N}(0, [\beta_0^{[t]}]^{-1})$ and $z_l^{[t]} \sim \text{Bern}(\pi_l^{[t]})$

- A small value of $\pi$ tends to generate a zero entry.
Structure-Aware Bayesian Compressive Sensing

Discouraged pattern
Neutral pattern
Encouraged pattern

OMP
Proposed
Jammer Suppression in GPS Receiver

Input SNR = -16 dB (before despreading), JNR = 25 dB
Output SJNR = 1.08 dB -> 19.81 dB (after despreading)
Multi-Sensor Data-Dependent Kernel

- The simple averaging of TFDs and AFs over different antennas disfavors cross-terms and enhances auto-terms
  - Auto-terms have same phase across all sensors
  - Cross-term phase depends on contributing signals
  - Particular effective when different sampling patterns are adopted in each sensor

- Modified AOK based on averaged AF

\[
\max_{\psi} \int_{0}^{2\pi} \int_{0}^{\infty} |A_{\Sigma}(r, \psi) \Phi(r, \psi)|^2 \, r \, dr \, d\psi
\]

subject to \( \Phi(r, \psi) = \exp\left(-\frac{r^2}{2\sigma(\psi)}\right) \)

\[
\frac{1}{4\pi^2} \int_{0}^{2\pi} \int_{0}^{\infty} |\Phi(r, \psi)|^2 \, r \, dr \, d\psi = \frac{1}{4\pi^2} \int_{0}^{2\pi} \sigma(\psi) \, d\psi \leq \alpha
\]

where \( A_{\Sigma}(r, \psi) = \frac{1}{N} \sum_{p=1}^{N} A_p(r, \psi) \)
Simulation Results: Baseline

60% of missing samples (single sensor, SNR=10 dB)
Simulation Results: Effect of Reduced SNR

- (a) single sensor
- (b) four sensors with same missing pattern
- (c) four sensors with different missing patterns

(60% of missing samples, SNR=4 dB)

- With array again from 4 sensors (6 dB), the performance is similar to the baseline case for both cases.
Simulation Results: With More Missing Samples

(a) single sensor

(b) four sensors with same missing pattern

(c) four sensors with different missing patterns

(70% of missing samples, SNR=10 dB)

- Similar performance achieved with different missing patterns, whereas array gain does not improve the same pattern case.
Example of DOA Estimation

4 sensors, 50% of missing samples, SNR=0 dB

- **Pseudo WVD**
- **Reconstructed**
- **Sparse TF-MUSIC**
- **MUSIC**
Conclusion – Part II

- Missing data generates noise-like artifacts
- Data-dependent kernels are effective to suppress such artifacts
- For multi-sensor array, adaptive kernel should be obtained from the combined observation at all sensors
- Sparse reconstruction enables effective TF signature estimation from randomly sampled FM signals
- If applicable, sensors should use different sampling patterns to combat the effect of missing samples
- Sparse TF-MUSIC is developed with demonstrated effectiveness
Thank you!

Contact info: http://yiminzhang.com/pub_area.html#tf