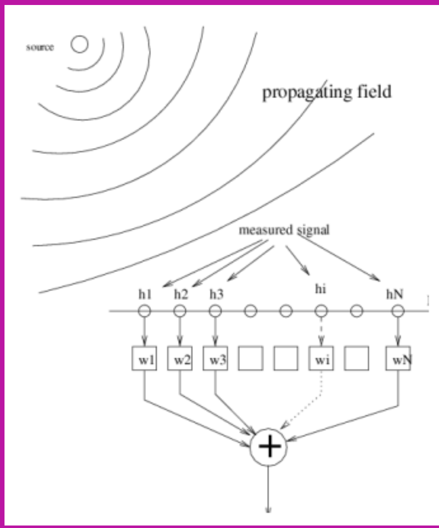
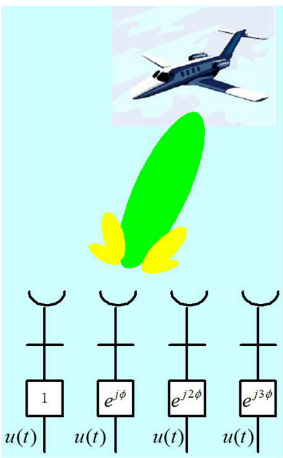
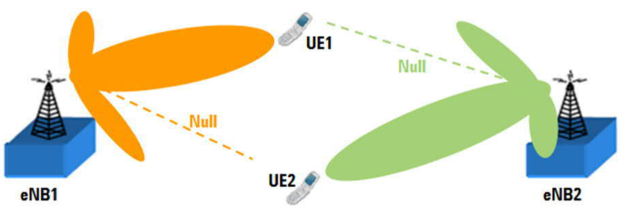
Multisensor Signal Processing

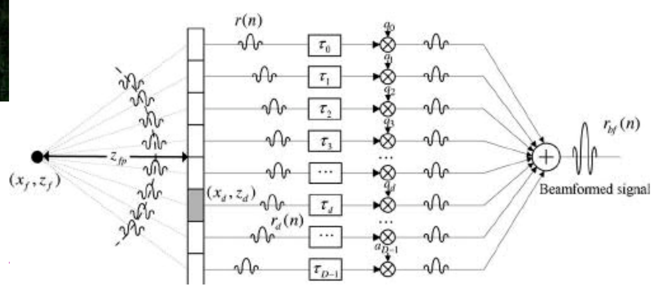


Aalto University
School of Engineering

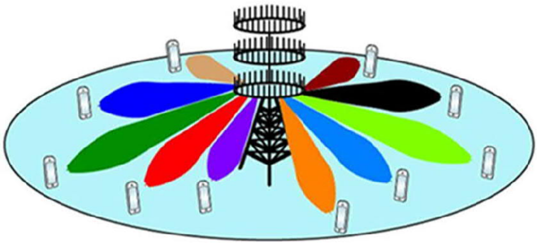
Sensor array signal processing

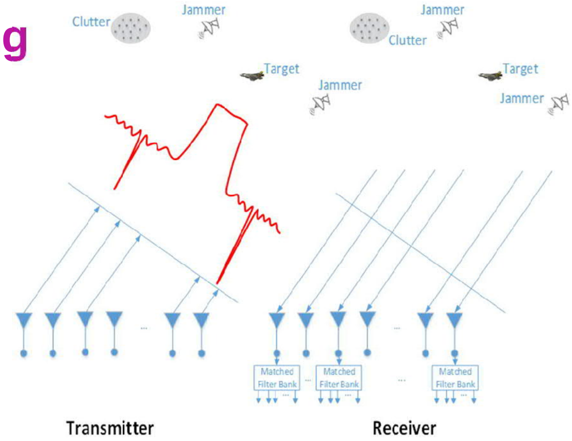


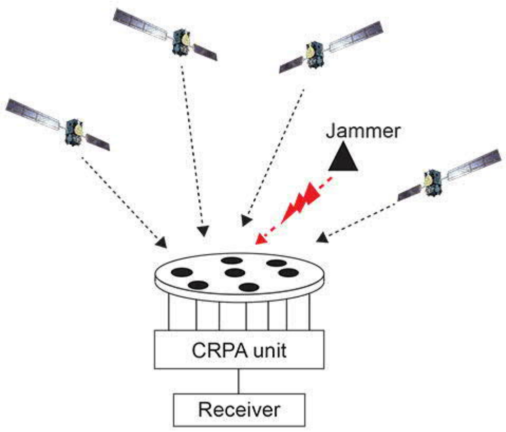




Sensor Array Signal Processing







Sensor array signal processing

- In array signal processing a group of sensors located at distinct spatial locations is employed.
- A propagating wavefield is measured with an array of M elements and a multichannel (M-channel) observation is formed.
- It is many times necessary to transmit/receive a signal to/from a certain direction and reject signals from other directions.
- This kind of spatial filtering is often called *beamforming*, and a system that performs such operation is referred to as a *beamformer*.

Sensor array signal processing

- Determining the location of the signal source or direction of arrival of the signal is often of interest
- Sampling may be non-uniform and arrays may be multidimensional.
- The key application areas include smart antennas in wireless communications and radar, channel sounding and propagation modeling, spatial sound processing, various electronic warfare applications, biomedical measurements such as EEG, MEG as well as sonar, seismology and radio astronomy.

Sensor array signal processing

- signal processing takes place in spatial or angular domain.
- In addition, time domain processing may be included and the resulting techniques are called Space-Time processing. This is necessary for broadband signals, e.g. in radar systems.
- Typically direction of arrival (DoA), number of signals and parameters of the waveform as well as channel parameters are estimated.

Sensor array signal processing

- If sensor array elements are uniformly spaced, the resulting signal processing methods are closely related to spectrum estimation and filtering in time domain.
- In this section an array processing signal model is presented
- Commonly used DoA estimation methods and their performance in statistical sense are described.
- It is many times necessary to receive a signal from a certain direction and reject signals from other directions.
- This kind of spatial filtering is often called *beamforming*, and a system that performs such operation is referred to as a *beamformer*.
- Adaptive beamforming algorithms are not considered here. See *Van Veen, B., Buckley, K., "Beamforming: A Versatile Approach to Spatial Filtering", IEEE ASSP Magazine, pp. 4-24, April 1988* for additional information.

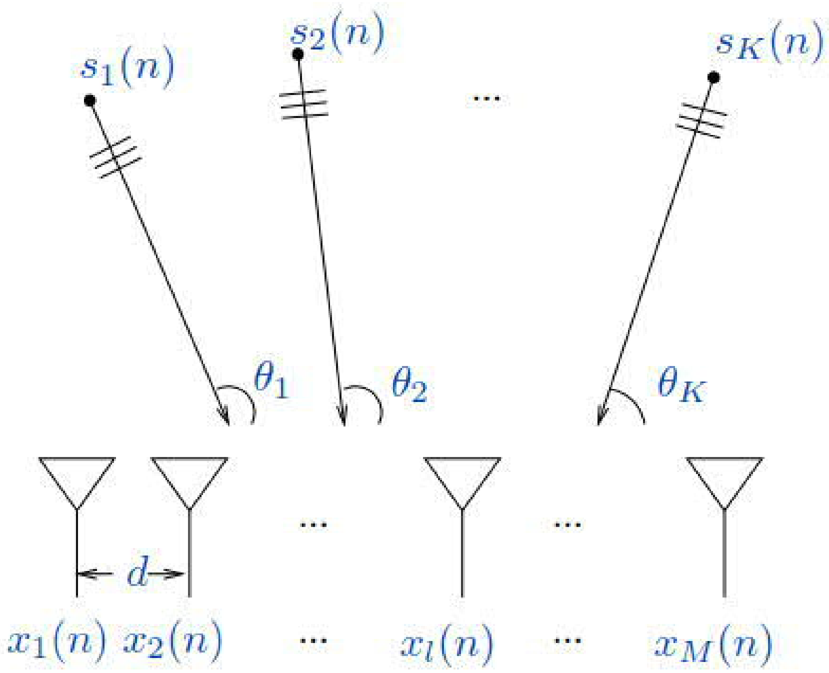
Sensor array signal processing

- Forming beams indicates to transmitting energy and sending signals, but it can be used for either transmission or reception.
- In radar, for example, interference usually occupies same frequency band as the signal of interest, so temporal filtering cannot be used to remove the harmful interference.
- The interference is usually radiating from different location than the desired signal, so beamforming can be exploited for filtering it.
- Other important properties are discrimination capability (resolution, aperture), adaptivity and capability of dealing with coherent (perfectly correlated) sources.

Array processing signal model

- Commonly used assumptions include
  - Narrowband assumption: bandwidth of the signal is very small compared to the center frequency
  - Sources are point emitters
  - Medium is homogeneous
  - Far field assumption: a plane wave is observed at the array and the propagation is characterized by pure delay
  - Sensors are spaced less than $\lambda/2$ (λ is wavelength) from each other to avoid spatial aliasing.
- DoA characterizes the source location if the above assumptions hold
- In wideband array processing, filterbanks and narrowband processing or space-time structures are needed.
- Notation: Let there be K signals present at an array of M sensors, $K < M$.

Array processing signal model



Array processing signal model

- ω_c is the center frequency, $g_k(\theta)$ represents the sensitivity of the kth sensor to the DOA θ and $\tau_k(\theta)$ is the time delay of the signal coming from DOA θ at the kth sensor relative to some reference point (typically the first sensor element).
- The positions and transfer characteristics of the array elements are assumed to be known.
- The collection of steering vectors over the parameter space of interest, Θ, is the called *array manifold*, $\mathcal{A}$,

$$\mathcal{A} = \{\mathbf{a}(\theta) \mid \theta \in \Theta\}.$$

- For K distinct DOAs $\theta_1, \ldots, \theta_K$ and the corresponding steering vectors are linearly independent
- The functions $\{g_k(\theta)\}_{k=1}^{M}$ depend on the type of sensors being used.

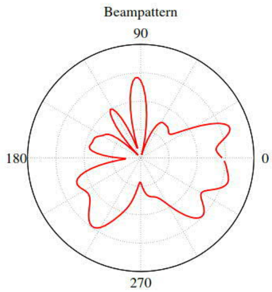
Array processing signal model

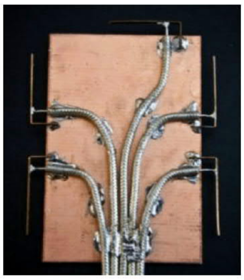
- ω_c is the center frequency, $g_k(\theta)$ represents the sensitivity of the kth sensor to the DOA θ and $\tau_k(\theta)$ is the time delay of the signal coming from DOA θ at the kth sensor relative to some reference point (typically the first sensor element).
- The positions and transfer characteristics of the array elements are assumed to be known.
- The collection of steering vectors over the parameter space of interest, Θ, is the called *array manifold*, $\mathcal{A}$,

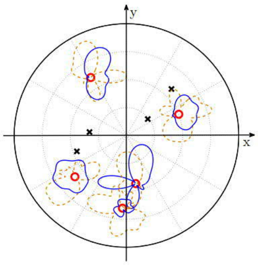
$$\mathcal{A} = \{\mathbf{a}(\theta) \mid \theta \in \Theta\}.$$

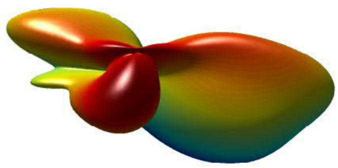
- For $K+1$ distinct DOAs $\theta_1, \ldots, \theta_K, \theta_{K+1}$ the corresponding steering vectors are linearly independent
- The functions $\{g_k(\theta)\}_{k=1}^{M}$ depend on the type of sensors being used.

Array processing signal model









Array processing signal model

- Often the sensors may be considered omnidirectional and identical, hence

$$\mathbf{a}(\theta) = [1, e^{-j\omega_c \tau_2(\theta)}, \ldots, e^{-j\omega_c \tau_M(\theta)}]^T.$$

- Let us have an array of M identical sensors uniformly spaced on a line, i.e., Uniform Linear Array (ULA)
- Now $\Theta = [0, \pi]$ and $\mathbf{a}(\theta)$ is given by

$$\mathbf{a}(\theta) = [1, e^{2\pi j (d/\lambda)\cos(\theta)}, \ldots, e^{2\pi j (M-1)(d/\lambda)\cos(\theta)}]^T$$

  where d denotes the element spacing and λ denotes the wavelength.
- In order to avoid spatial aliasing $d \leq \lambda/2$.
- There are some special nonuniform array structures (sparse arrays) where element spacing is designed to give high aperture and additional degrees of freedom with fewer elements.

Array processing signal model

- Matrix A is then a Vandermonde matrix, i.e.

$$A =$$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{2\pi j(d/\lambda)\cos(\theta_1)} & e^{2\pi j(d/\lambda)\cos(\theta_2)} & \cdots & e^{2\pi j(d/\lambda)\cos(\theta_K)} \\ e^{2\pi j2(d/\lambda)\cos(\theta_1)} & e^{2\pi 2j(d/\lambda)\cos(\theta_2)} & \cdots & e^{2\pi 2j(d/\lambda)\cos(\theta_K)} \\ \vdots & \cdots & \cdots & \vdots \\ e^{2\pi j(M-1)(d/\lambda)\cos(\theta_1)} & e^{2\pi j(M-1)(d/\lambda)\cos(\theta_2)} & \cdots & e^{2\pi j(M-1)(d/\lambda)\cos(\theta_K)} \end{bmatrix}$$

- If the angle is against broadside of the array, we need to have $-\sin(\theta)$ above.

EX: Uniform Circular Array (UCA) model

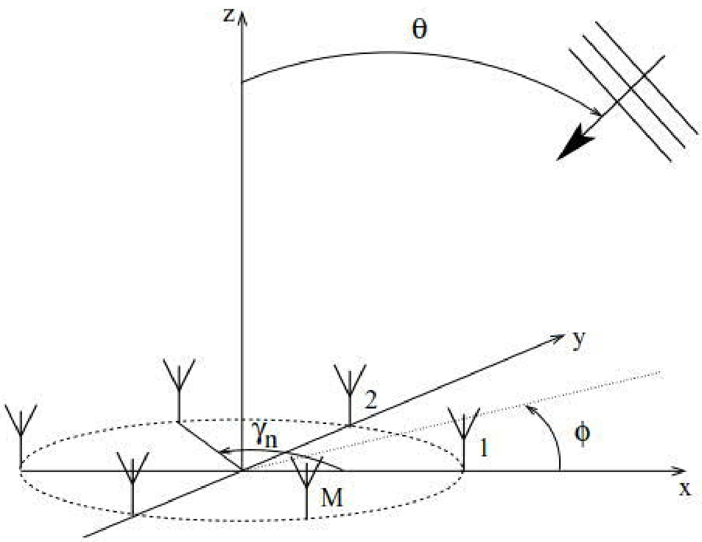
- In case of 2-D arrays such as uniform circular array (UCA), both elevation and azimuth angles may be estimated.
- In UCA, the antenna elements are uniformly spaced on a circle of diameter $2r$.
- The azimuth and elevation angles are denoted by ϕ_l and θ_l.
- The $M \times 1$ element-space UCA steering vector gets then form as

$$\mathbf{a}(\theta,\phi) = \begin{pmatrix} e^{j\omega\frac{r}{c}\sin\theta\cos(\phi-\gamma_0)} \\ e^{j\omega\frac{r}{c}\sin\theta\cos(\phi-\gamma_1)} \\ \vdots \\ e^{j\omega\frac{r}{c}\sin\theta\cos(\phi-\gamma_{M-1})} \end{pmatrix},$$

where $\omega = 2\pi f$ is the angular frequency, r is the radius of the array, $\gamma_i = \frac{2\pi i}{M}$ is the angular position of the element (counted in counterclock wise manner from x-axis) and c is the speed of light.

EX: Uniform Circular Array (UCA) model

- Uniform Circular Array (UCA) with M elements. Elevation angle θ and azimuth angle ϕ may be estimated simultaneously.



Array processing algorithms

- The following methods are studied: classical Beamformer, Capon's Minimum Variance Distortionless Response (MVDR) method, subspace methods (MUSIC, ESPRIT) as well as Deterministic and Stochastic Maximum Likelihood methods. Subspace fitting is considered as well.
- Similar signal methods are often used for time-domain processing and time-series data. Beamforming is directly related to FIR-filtering, periodograms and using windowing functions in spectrum estimation. MVDR has also many applications in filtering and communications SP.
- High resolution methods (MUSIC, ESPRIT) can be used for frequency estimation and resolving closely spaced frequencies based on autocorrelation matrix and its eigenvalue decomposition as well as delay estimation.

Beamforming: spatial filtering

- Let us denote the observations acquired by the array by $\mathbf{x}(n)$, the vector of spatial filter coefficients by $\mathbf{w}$ and the output of the spatial filter $y(n)$, i.e.

$$y(n) = \mathbf{w}^H \mathbf{x}(n)$$

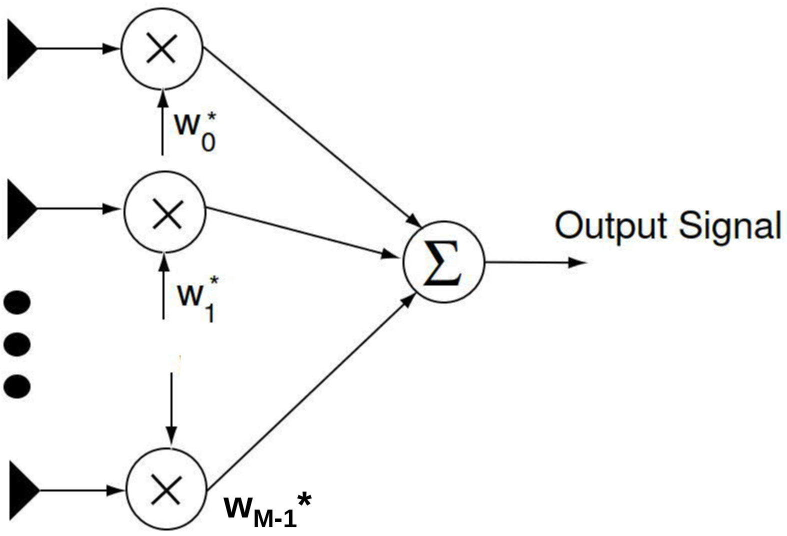
- The power of the spatially filtered signal is

$$E[|y(n)|^2] = \mathbf{w}^H \Sigma \mathbf{w}$$

  and it should exhibit a peak in the direction where an emitter is transmitting. $\Sigma = E[\mathbf{x}(n)\mathbf{x}^H(n)]$ is the covariance matrix of the observations.

- In Beamforming we want to find optimal filter coefficients, typically to maximize SNR.

Beamforming: spatial filtering



Beamforming: spatial filtering

- Beamformer is a direct extension of periodogram to spatial processing (averaging spatial periodograms)
- In beamforming the the array is steered in one direction at the time and the power is measured.
- The beamformer is designed so that the output power $E[|y(n)|^2]$ from given direction θ is maximized:

$$E[\mathbf{w}^H\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}] = \mathbf{w}^H\Sigma\mathbf{w}$$

- Steering to the DoA is done by choosing by weighting ($\mathbf{w}$) the sensor outputs so that output SNR is maximized subject to $\mathbf{w}^H\mathbf{w} = 1$. The constraint is needed to avoid trivial solution.
- Assume that there is one WSS signal present at angle θ_1. The array output is then

$$\mathbf{x}(n) = \mathbf{a}(\theta_1)s_1(n) + \mathbf{v}(n).$$

Beamforming: spatial filtering

- The power of $y(n)$ is given as

$$E[y(n)y^*(n)] = \mathbf{w}^H\Sigma\mathbf{w} = \sigma_s^2\mathbf{w}^H\mathbf{a}(\theta_1)\mathbf{a}^H(\theta_1)\mathbf{w} + \sigma_v^2\mathbf{w}^H\mathbf{w},$$

- The problem of maximizing the output SNR may now be formulated as

$$\max_{\mathbf{w}} \{\mathbf{w}^H\mathbf{a}(\theta_1)\mathbf{a}^H(\theta_1)\mathbf{w}\} = \max_{\mathbf{w}} ||\mathbf{w}^H\mathbf{a}(\theta_1)||^2, \text{ subject to } \mathbf{w}^H\mathbf{w} = 1,$$

- From Cauchy-Schwartz and the constraint $\mathbf{w}^H\mathbf{w} = 1$

$$||\mathbf{w}^H\mathbf{a}(\theta_1)||^2 \leq ||\mathbf{w}||^2||\mathbf{a}(\theta_1)||^2 = ||\mathbf{a}(\theta)||^2.$$

Beamforming: spatial filtering

- The optimal weighting vector is therefore given as

$$\mathbf{w}_{BF} = \frac{\mathbf{a}(\theta_1)}{\sqrt{\mathbf{a}^H(\theta_1)\mathbf{a}(\theta_1)}}.$$

- The denominator can be considered to be a normalization factor of the steering vectors $\mathbf{a}^H(\theta)$
- This can be considered to be a spatial filter matched to the impinging signal.

Beamforming: spatial filtering

- Using the optimal weighting, spatial spectrum is given by

$$V_{BF}(\theta_1) = \frac{\mathbf{a}^H(\theta_1)\Sigma\mathbf{a}(\theta_1)}{\mathbf{a}^H(\theta_1)\mathbf{a}(\theta_1)}.$$

- This is a spatial extension of the Periodogram method in Spectrum Estimation.
- In general, the optimal coefficients are

$$\mathbf{w}_{BF} = \frac{\mathbf{a}(\theta)}{\sqrt{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}}.$$

- The Beamforming estimates of the K DoA's are the K highest peaks of

$$\mathbf{a}^H(\theta)\Sigma\mathbf{a}(\theta).$$

Beamforming: spatial filtering

- array covariance matrix Σ is not typically known and it may be estimated by

$$\hat{\Sigma} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}(n)\mathbf{x}^H(n).$$

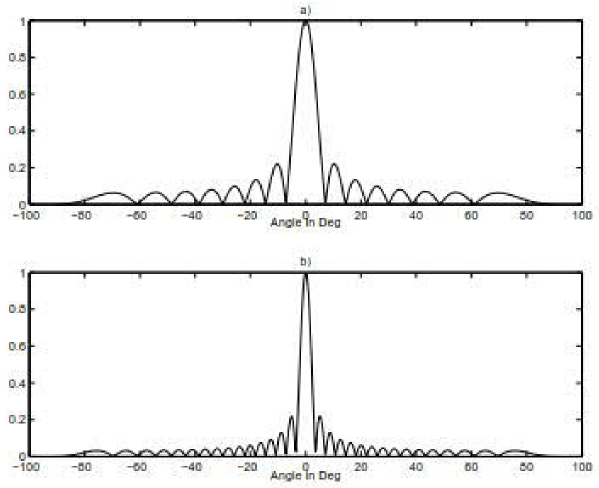
  using the obtained snapshots.
- Maximum likelihood estimator if $K = 1$ and consequently consistent.
- When $K > 1$ beamforming DOA estimates are the angles of the K highest peaks in spatial spectrum. These estimates are not consistent then and asymptotic bias may be large if sources correlated or closely spaced.

Beamforming resolution

- The Beamformer can not resolve closely spaced DoA's. For a ULA of M sensors, the beamforming resolution limit is approximately $\frac{\lambda}{Md}$ (i.e. wavelength / array length).
- Beamwidth of the spatial filter is inversely proportional to array aperture ($\approx$ array length in wavelengths)
- For example, for a ULA of 6 sensors of $\lambda/2$ inter-element spacing, the resolution limit equals $1/3$ rad $\approx 19°$.
- For ULA's, the beamformer output is obtained through normalized spatial DFT of $x(n)$
- Different windowing functions may be used similarly to time-domain spectrum estimation.

Beamforming resolution

- Beamformer resolution depends on the number of elements (example: 16 and 64 elements below).



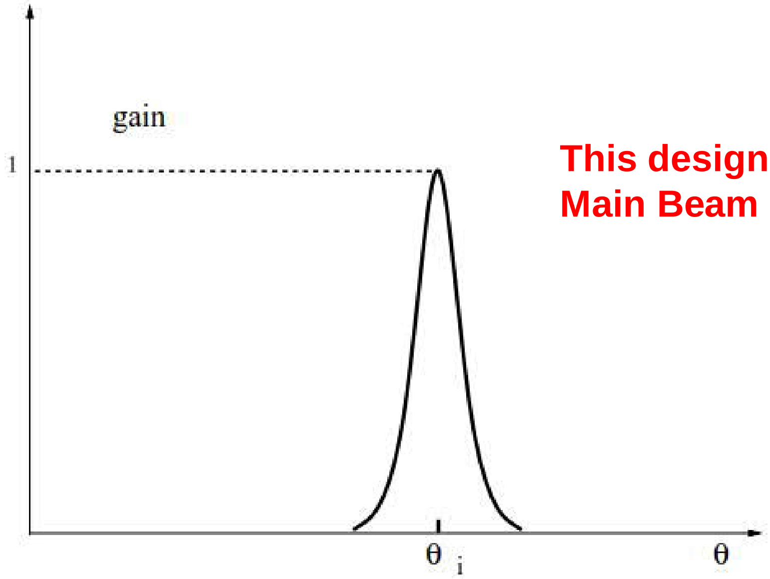
Array aperture: Size of array in wavelenghts

Minimum Variance Distortionless Response (MVDR) or Capon Beamformer

- MVDR (Minimum Variance Distortionless Response) was derived to improve the resolution of the beamformer.
- In the presence of multiple sources the power measured by the spatial spectrum is not only due to the power of the source at that direction, but also to power of other sources in other directions.
- MVDR minimizes the total output power of $\mathbf{y}(n) = \mathbf{w}^H \mathbf{x}(n)$ while maintaining the gain along the look direction θ constant (unity, hence distortionless or unbiased)
- The spatial filter response is forced to be narrowband. Consequently, the output power is due mostly to the power from the look direction.
- Therefore, it tends to steer nulls towards interferers.
- Let $\mathbf{x}(n)$ be WSS and $\Sigma = E[\mathbf{x}(n)\mathbf{x}^H(n)]$.

MVDR Beamformer

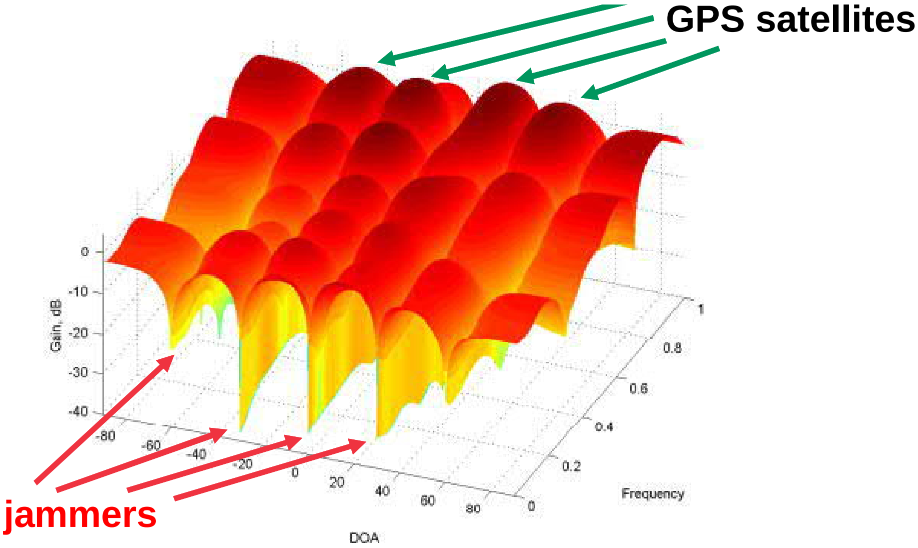
- Spatial filtering in MVDR: gain in direction θ_i is constrained to unity and total output power is minimized.



This design tries to make the Main Beam as narrow as possible

EX: MVDR Beamformer in GPS antijamming

- Gain in direction of satellites, nulls towards broadband and narrowband jammers. Implemented using space-time Householder multistage wiener filters (MSWF)



MVDR Beamformer

- MVDR problem is then formulated

$$\min_{\mathbf{w}} \mathbf{w}^H \Sigma \mathbf{w} \text{ subject to } \mathbf{w}^H \mathbf{a}(\theta) = 1.$$

- The following Lagrange expression is minimized

$$J(\mathbf{w}) = \mathbf{w}^H \Sigma \mathbf{w} + \nu(1 - \mathbf{w}^H \mathbf{a}(\theta))$$

wrt. $\mathbf{w}$, and ν denotes a Lagrange multiplier.

- The solution is found by

$$\mathbf{w}_{MVDR} = \frac{\Sigma^{-1}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\Sigma^{-1}\mathbf{a}(\theta)}.$$

- In practice, an estimate for $\mathbf{w}_{MVDR}$ is formed from N snapshots $\mathbf{x}(1), \ldots, \mathbf{x}(N)$.

MVDR Beamformer

- Using sample estimate of covariance $\hat{\Sigma}$:

$$\hat{\mathbf{w}}_{MVDR} = \frac{\hat{\Sigma}^{-1}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\hat{\Sigma}^{-1}\mathbf{a}(\theta)}.$$

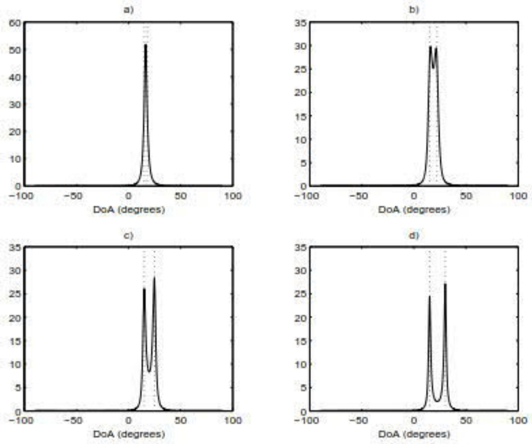
- The DoA's are the K highest peaks in the spatial spectrum

$$V_{MVDR}(\theta) = \frac{1}{\mathbf{a}(\theta)^H \Sigma^{-1}\mathbf{a}(\theta)}.$$

- A significant improvement in the resolution is obtained. However, it depends on M and on the SNR.
- The performance obviously deteriorates if signals are correlated due to multipath.
- Method is also sensitive to errors in array manifold
- In adaptive beamforming a recursive formula may be used

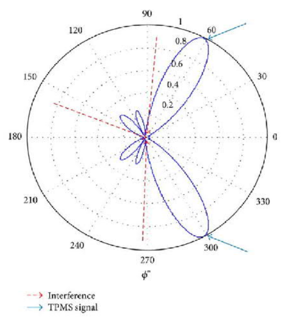
MVDR Beamformer resolution

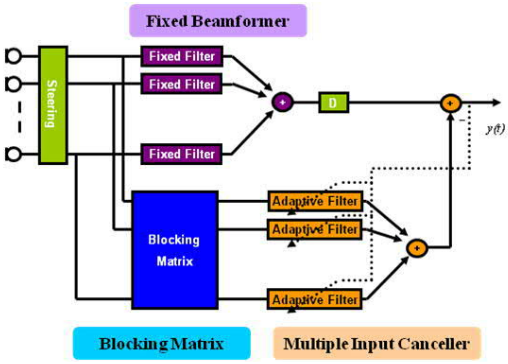
- Resolving close spaced sources using MVDR, difference in DoA is (a) 3, (b) 7, (c) 10 and (d) 15 degrees. SNR 15 dB, 8-element array, 256 snapshots.



MVDR Beamformer

- Estimates are not consistent for $K > 1$
- So called Generalized Sidelobe Canceller (GSC) solves the same problem by reformulating it into unconstrained one. This allows for simpler implementation and the use of well known adaptive filter structures.





Generalized Sidelobe Canceller (GSC)

- GSC provides a mechanism changing the constrained minimization problem in MVDR to an unconstrained one.
- Let the constraint be of form $\mathbf{a}^H(\theta)\mathbf{w} = g$, where g is a constant (typically $g = 1$).
- The single constraint may be generalized to several constraints.
- For example, if there is a fixed interferer in known direction, the gain in that direction may be forced to zero.
- In case of two constraints, this would be expressed as

$$\begin{bmatrix} \mathbf{a}^H(\theta_1) \\ \mathbf{a}^H(\theta_2) \end{bmatrix} \mathbf{w} = \begin{bmatrix} g \\ 0 \end{bmatrix}$$

- The L constraints may be written in the form $C^H\mathbf{w} = \mathbf{f}$ where C is constraint matrix and $\mathbf{f}$ response vector. Constraints are assumed to be linearly independent.

Generalized Sidelobe Canceller (GSC)

- Suppose we decompose the weight vector $\mathbf{w}$ into two orthogonal components $\mathbf{w}_0$ and $\mathbf{v}$ that lie in the range (column) and null space of matrix C.
- The range and null space span the entire space so the decomposition may be used to represent any $\mathbf{w}$
- Since $C^H\mathbf{v} = 0$, we must have

$$\mathbf{w}_0 = C(C^H C)^{-1}\mathbf{f}$$

  if $\mathbf{w}$ satisfies the constraints
- Let C_n form a basis to null space of the matrix C. The vector $\mathbf{v}$ is a linear combination of the columns of (blocking) matrix C_n, (i.e. $\mathbf{v} = C_n\mathbf{w}_M$)
- C_n may be obtained from C by using SVD or QR decomposition, for example.

Generalized Sidelobe Canceller (GSC)

- The weight vector is

$$\mathbf{w} = \mathbf{w}_0 - C_n \mathbf{w}_M$$

- The choice for $\mathbf{w}_0$ and C_n implies that $\mathbf{w}$ satisfies the constraints independent of $\mathbf{w}_M$ and reduces the MVDR problem to an unconstrained one

$$\min_{\mathbf{w}_M} \; [\mathbf{w}_0 - C_n \mathbf{w}_M]^H \Sigma [\mathbf{w}_0 - C_n \mathbf{w}_M]$$

- The solution is

$$\mathbf{w}_M = (C_n^H \Sigma C_n)^{-1} C_n^H \Sigma \mathbf{w}_0$$

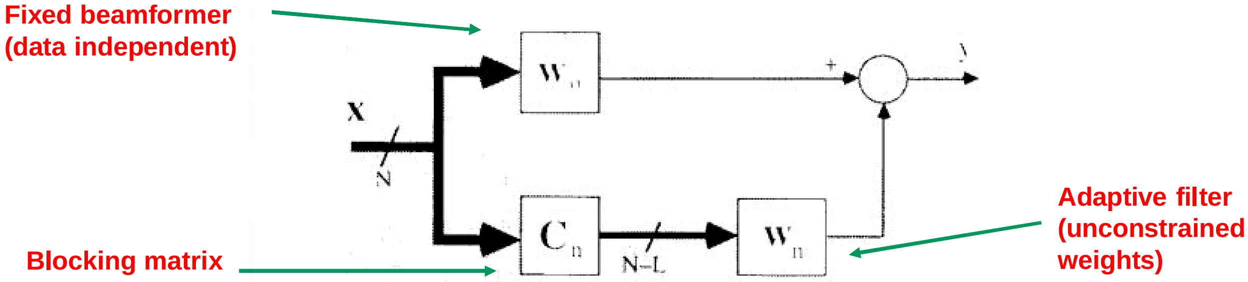
- Jargon: $\mathbf{w}_0$ is fixed beamformer, C_n is called blocking matrix (will block the DOA's given in constraints and those signals are processed by $\mathbf{w}_0$) and unconstrained weights $\mathbf{w}_M$ may be computed using standard adaptive filter

Generalized Sidelobe Canceller (GSC)

- By setting the output of the fixed beamformer $d(n) = \boldsymbol{w}_0^H \boldsymbol{x}(n)$ as the desired signal, and $\boldsymbol{u}(n) = C_n^H \boldsymbol{x}(n)$, and by defining error signal $e(n) = d(n) - \boldsymbol{w}_M^H \boldsymbol{u}(n)$, $R_u = Cov(u)$ we get the following LMS update for the beamformer coefficients:

$$\boldsymbol{w}_M(n+1) = \boldsymbol{w}_M(n) + \mu \boldsymbol{u}(n) e^*(n)$$

- Filter is initialized by $\boldsymbol{w}(0) = 0$, $e(0) = d(0)$ and $0 \leq \mu \leq \frac{1}{Trace(R_u)}$



GSC using LMS or Recursive Least Square (RLS)

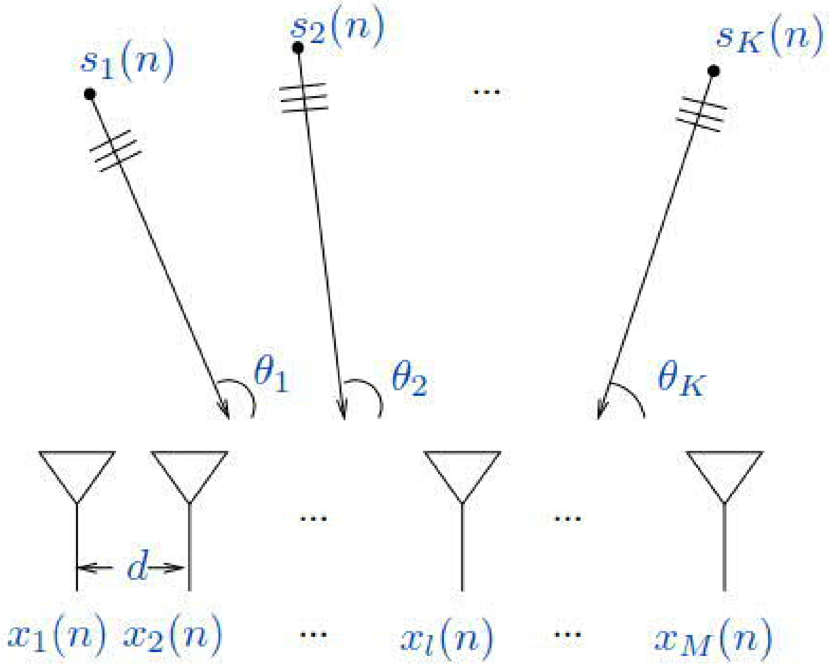
- LMS and RLS weight adaptation algorithms. $\mathbf{u}(n)$ is the data signal and $d(n)$ is the desired signal.

Algorithm	LMS	RLS
Initialization	$\mathbf{w}(0) = \mathbf{0}$ $e(0) = d(0)$ $0 < \mu < \frac{1}{Trace[\mathbf{R}_u]}$	$\mathbf{w}(0) = \mathbf{0}$ $\mathbf{P}(0) = \delta \mathbf{I}$ δ small, $\mathbf{I}$ identity matrix
Update Equations	$\mathbf{w}(n) = \mathbf{w}(n-1)$ $+ \mu \mathbf{u}(n-1) e^*(n-1)$ $e(n) = d(n) - \mathbf{w}^H(n)\mathbf{u}(n)$	$\mathbf{q}(n) = \mathbf{P}(n-1)\mathbf{u}(n)$ $\mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{q}(n)}{1+\lambda^{-1}\mathbf{u}^H(n)\mathbf{q}(n)}$ $\alpha(n) = d(n) - \mathbf{w}^H(n-1)\mathbf{u}(n)$ $\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{k}(n)\alpha^*(n)$ $\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{q}^H(n)$

Subspace methods in array processing

- We will study subspace DoA estimation techniques
  - MUSIC (MUltiple SIgnal Classification)
  - TLS-ESPRIT (Total LS-Estimation of Signal Parameters via Rotational Invariance Techniques)
- Such techniques are high-resolution techniques that can resolve arbitrarily closely spaced angles of arrival.
- The groundwork was done by V.F. Pisarenko (Pisarenko Harmonic Decomposition)
- These techniques are related to eigen-decomposition of the array covariance matrix.
- They decompose the observation vector space in a subspace associated with the signal and a subspace associated with noise.

Array processing signal model



Subspace methods in array processing

- Recall the signal model

$$\mathbf{x}(n) = A(\theta)\mathbf{s}(n) + \mathbf{v}(n).$$

- Assume that we have K signals present and the $K \times K$ signal covariance matrix $\Sigma_s = E\{\mathbf{s}(n)\mathbf{s}^H(n)\}$ is of rank K.
- If signals are highly correlated, some preprocessing steps are required such as spatial smoothing.
- The noise was assumed to be spatially and temporally white and independent from the signals.
- The array $M \times M$ covariance matrix of $\mathbf{x}(n)$ is

$$\Sigma = E[\mathbf{x}(n)\mathbf{x}^H(n)] = A\Sigma_s A^H + \sigma^2 I$$

where σ^2 is the noise variance.

Subspace methods in array processing

- The eigevaluedecomposition of Σ is

$$\Sigma = U\Lambda U^H$$

- Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K > \lambda_{K+1} = \cdots = \lambda_M = \sigma^2$ be the eigenvalues of Σ (diagonal elements of Λ).
- The $M - K$ smallest eigenvalues of Σ are equal to σ^2 and the corresponding eigenvectors are orthogonal to the columns of A.
- These eigenvectors U_n span the *noise subspace*.
- The eigenvectors U_s corresponding to the K largest eigenvalues span the *signal subspace*.
- Consequently, the covariance matrix Σ may also be decomposed as

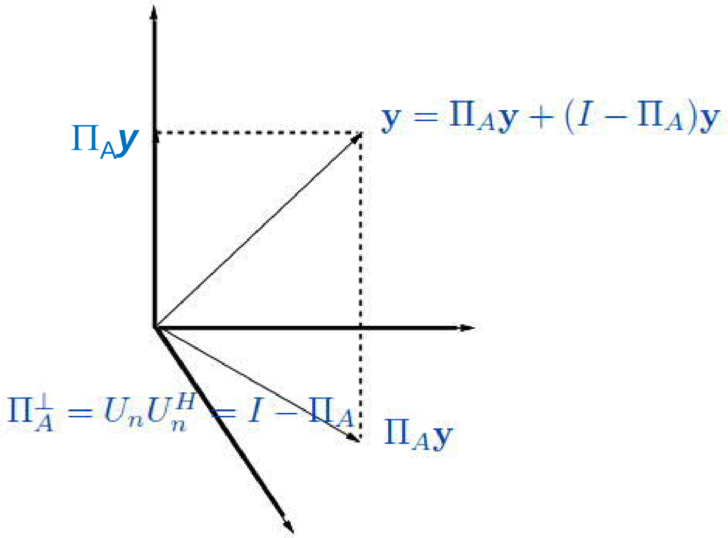
$$\Sigma = [U_s \; U_n] \; \mathrm{diag}\{\lambda_1, \ldots, \lambda_M\} \; [U_s \; U_n]^H.$$

Subspace methods in array processing

- The projection matrix to the signal subspace is $\Pi_A = U_s(U_s^H U_s)^{-1} U_s^H = U_s U_s^H$.
- The columns of A also span the signal subspace. Hence, $\Pi_A = AA^\dagger$, where $A^\dagger = (A^H A)^{-1} A^H$.
- The projection matrix to the orthogonal noise subspace is $\Pi_A^\perp = U_n U_n^H = I - \Pi_A$.
- Depending whether noise or signal subspace is used in the DoA estimation method, the methods are called noise subspace methods and signal subspace methods.
- MUSIC methods is a noise subspace method that exploits the fact that signals are orthogonal to the entire noise subspace.
- ESPRIT methods is a signal subspace method.
- Other methods include Weighted Subspace Fitting and Minimum norm method, for example.
- Maximum likelihood methods employ the projection matrices Π above.

Subspaces and projection matrices

$$\Pi_A = AA^\dagger = A(A^H A)^{-1} A^H = U_s U_s^H$$



MUSIC Subspace method

- MUSIC (MUltiple SIgnal Classification) by Schmidt 1979.
- MUSIC uses a covariance matrix of size $M > K + 1$ and computes its eigenvalues and eigenvectors.
- Eigenvalues form 2 groups: those corresponding to signal subspace and those corresponding to noise subspace. The eigenvalues corresponding to noise subspace are small and almost equal.
- If the number of signals K is not known a priori it can be found from the pattern of eigenvalues.
- For low SNR's the pattern may not be that obvious. In such cases criteria such as Akaike Information Criterion (AIC), Bayesian IC (BIC) or Minimum Description Length (MDL by Rissanen) may be used.
- The method uses projection of signal to the entire noise subspace.

MUSIC Subspace method

- Because of the orthogonality of the signal and noise subspace,
$$\mathbf{a}^H(\theta_i)U_nU_n^H\mathbf{a}(\theta_i) = 0$$
at the true DOAs $\theta_i, \quad i = 1, \ldots, K$. If $\theta \neq \theta_i$, the denominator is ≥ 0.
- Each of the signals (direction of arrival vectors) is orthogonal to the entire noise subspace, hence the projection goes to zero for true DoA's.
- When U_n is estimated using the matrix of the eigenvectors $\hat{U}_n$ corresponding to the $M-K$ smallest eigenvalues of the sample covariance matrix $\hat{\Sigma}$, the pseudo-spectrum
$$\hat{V}_M(\theta) = \frac{1}{\mathbf{a}^H(\theta)\hat{U}_n\hat{U}_n^H\mathbf{a}(\theta)}$$
will exhibit large peaks at the correct DOAs due to the orthogonality.

MUSIC Subspace method

- This expression holds for colored noise as well.
- The estimates of the DOAs are the K largest peaks in this pseudo-spectrum.
- The height of the peak gives an index of orthogonality and does not describe the signal power.
- MUSIC can resolve sources with arbitrary close DOAs.
- Maximum of $M-1$ DOAs can be estimated with an M element array.
- The estimates are consistent
- MUSIC does not achieve CRLB for finite M (Stoica and Nehorai).
- For uncorrelated sources, very good performance is achieved. For correlated sources the performance deteriorates.

MUSIC Subspace method

- Small sample sizes may yield poor results as well.
- In practise, resolution and variance depend on SNR, array size and number of snapshots.
- Typically at least $3 - 4\ dB$ SNR is required so that the low rank model works well.
- Subspace tracking (or root tracking in case of ULA's) algorithms may be used to reduce computation.
- Root-MUSIC may be applied to ULA's. It uses eigenfilter formulation and polynomial rooting.

Root-MUSIC method

- The denominator of the pseudospectrum expression is presented as a polynomial of z using eigenfilter formulation

$$U_i(z) = u_i(0) + u_i(1)z^{-1} + \cdots + u_i(N-1)z^{-(M-1)},$$

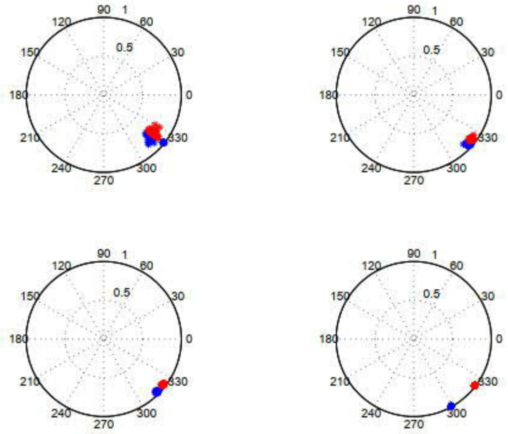
  where $u_i(k)$ are components of the eigenvector $\mathbf{u}_i$
- The expression $\mathbf{a}^H U_n U_n^H \mathbf{a}$ may now be rewritten

$$\sum_{i=K+1}^{M} \mathbf{a}^H \mathbf{u}_i \mathbf{u}_i^H \mathbf{a} = \sum_{i=K+1}^{M} U_i(z) U_i^*(1/z^*)$$

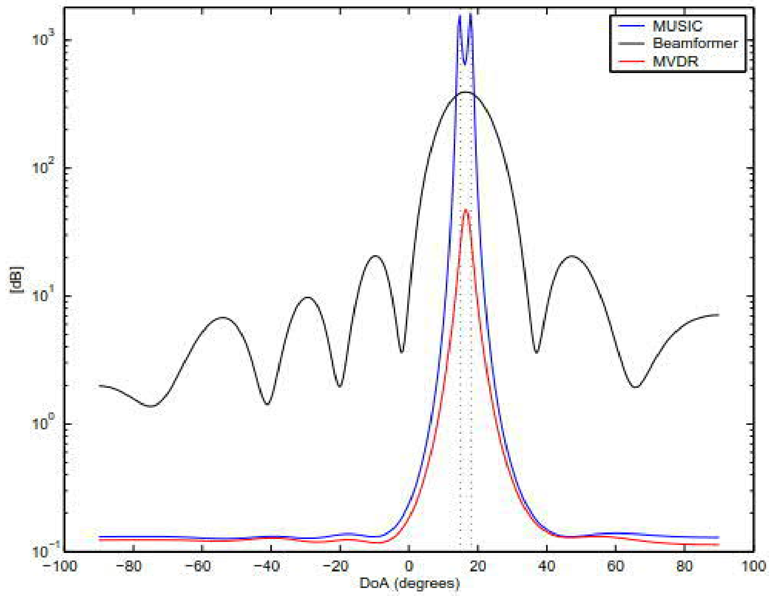
  and it is set to zero and its roots are found.
- The AoA is associated with the phase angle of the K roots on (close to) the unit circle. The remaining zeros are spurious.

Root-MUSIC method

- Then exhaustive search for steering vectors (orthogonality) may be avoided with Root-MUSIC.
- Roots in case of 2 sources. Difference in AoA (from upper left corner): 1, 2, 3 and 20 degrees.



High resolution in direction of arrival estimation



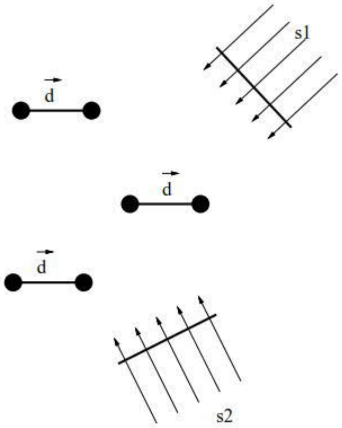
ESPRIT method

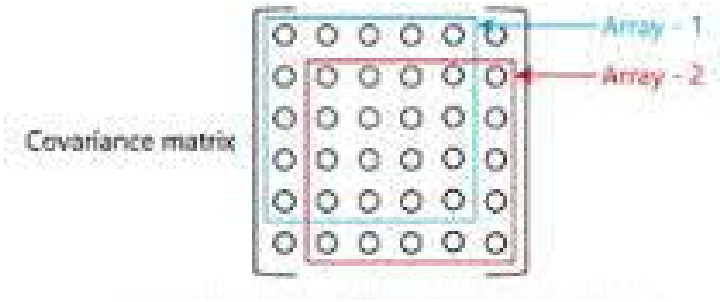
- ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique)
  - "son" of MUSIC
  - The original technique by Paulraj, Roy and Kailath 1985.
  - an improved version by Roy and Kailath 1989.
  - Can be used only for special array configurations
- Significantly lower computational complexity than MUSIC. No exhaustive search for steering vectors needed.
- The original ESPRIT and the version using Total LS (TLS-ESPRIT) are described here.
- Matrix A is not necessarily precisely known.
- ESPRIT assumes that the sensor array can be decomposed into two identical subarrays separated by some fixed displacement vector.

ESPRIT method

- Subarrays do not need to calibrated, just identical (displacement between subarrays accurate)
- The subarrays may be overlapping, e.g., in case of ULA first subarray elements $1, \ldots, M-1$ and second subarray elements $2, \ldots, M$.
- Denote the dimension of the twin subarrays by P. Let J_1 be the $P \times M$ matrix that selects the left subarray from the array output vector.
- Let J_2 denote the corresponding matrix for right subarray.
- In the ULA example above, $P = M - 1$ and the selection matrices J_1 and J_2 are given as $J_1 = [I_{M-1}\ \mathbf{0}]$ and $J_2 = [\mathbf{0}\ I_{M-1}]$.
- Let $A_1 = J_1 A$ and let $A_2 = J_2 A$ and A_1, A_2 be of full column rank

ESPRIT method





Covariance matrix for 2 overlapping arrays of 5 elements, total of 6 array elements

ESPRIT method

- ESPRIT exploits the following shift invariance property

$$J_1 A \Phi = J_2 A$$

i.e.

$$A_2 = A_1 \Phi,$$

where $\Phi = diag(\phi_1, \phi_2, ..., \phi_K)$ is a diagonal matrix and

$$\phi_i = \exp\{j 2\pi (d/\lambda) \cos(\theta_i)\}, \quad i = 1, \ldots, K.$$

where $d \leq \lambda/2$ is the displacement and θ_i are DoA's.

- Angles of arrival θ_i are obtained from ϕ_i's by

$$\theta_i = acos(\frac{\lambda}{2\pi d} \text{arg}\{\phi_i\}).$$

ESPRIT method

- Typically, the angle is measured against broadside of the array then $\Phi = diag(e^{-j\frac{2\pi}{\lambda}d\ sin\theta_i}), i = 1, ..., K$
- Matrix A and signal subspace eigenvectors U_s span the same column space (Range space of U_s is the range space of array steering matrix A)
- Consequently, there exist a nonsingular $K \times K$ matrix C such that $U_sC = A$

$$\begin{cases} J_1U_sC = J_1A \\ J_2U_sC = J_2A. \end{cases}$$

- Substituting the above to the expression on shift invariance property

$$J_1U_sC\Phi = J_2U_sC \quad \text{we get} \quad J_2U_s = J_1U_s\Psi,$$

where $\Psi = C\Phi C^{-1}$.

ESPRIT method

- Φ is a diagonal, hence $C\Phi C^{-1}$ is in a form of eigenvalue decomposition.
- A result from linear algebra says that that Ψ and Φ have the same eigenvalues.
- Now by finding Ψ we get Φ and using ϕ_i we find θ_i.
- Estimate $\hat{U}_s$ of U_s is formed from array outputs and the following approximate invariance equation

$$J_2U_s \approx J_1U_s\Psi,$$

may be solved using the Least Square or TLS approach

$$J_2\hat{U}_s = J_1\hat{U}_s\Psi,$$

because the equality does not hold exactly.

ESPRIT method

- The corresponding methods are coined LS-ESPRIT and TLS-ESPRIT
- TLS is a natural choice because the error can be assumed to be identically distributed on LHS and RHS of the above equation.
- The resulting Ψ may be expressed in terms of eigendecomposition $\Psi = C\Phi C^{-1}$, where eigenvalues $\Phi = diag(\phi_i)$, $i = 1, ..., K$
- DOA's θ_i are obtained from ϕ_i's by

$$\theta_i = acos(\frac{\lambda}{2\pi d}\text{arg}\{\phi_i\}).$$

- Correspondingly, if angle against broadside is used spatial frequencies

$$\mu_i = -\frac{2\pi}{\lambda} d\ sin(\theta_i) = arg(\phi_i)$$

ESPRIT method

- The DOA is:

$$\theta_i = -\frac{\lambda}{2\pi d}\ arcsin(\mu_i), \quad 1 \le i \le K.$$

- The estimates are consistent
- The asymptotic variance is larger than that of MUSIC.
- Can not be used in the presence of coherent sources without appropriate preprocessing (spatial smoothing)

Weighted Subspace Fitting (WSF) method

- A method originally developed by Viberg and Ottersten.
- The method links the subspace methods and maximum likelihood methods. In fact, asymptotically optimal performance may be achieved similarly to Stochastic ML method.
- The rank of the signal covariance matrix Σ_s $K' \leq K$
- the eigenvalues of Σ are $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{K'} > \lambda_{K'+1} = \ldots = \lambda_M = \sigma^2$
- The corresponding eigenvectors are $\mathbf{u}_1, \ldots, \mathbf{u}_M$.
- It is assumed that $K < (M + K')/2$ to ensure identifiability.
- Let us denote $\Lambda_s = diag[\lambda_1, \ldots, \lambda_{K'}]$, $U_s = [\mathbf{u}_1, \ldots, \mathbf{u}_{K'}]$, $U_n = [\mathbf{u}_{K'+1}, \ldots, \mathbf{u}_M]$

Weighted Subspace Fitting (WSF) method

- If Σ_s is of full rank ($K' = K$), then A and U_s span the same column space.
- In case $K' < K$, the space spanned by the columns of U_s is contained in the space spanned by the columns of A.
- Consequently, there has to be a unique $K \times K'$ transformation matrix $\mathcal{T}$ such that

$$U_s = A\mathcal{T}$$

- In subspace fitting we want to find the best transform such that

$$[\hat{\theta}, \hat{\mathcal{T}}] = \arg\min_{\theta, \mathcal{T}} ||U_s - A\mathcal{T}||_F^2$$

is optimized.

Weighted Subspace Fitting (WSF) method

- A solution for $\mathcal{T}$ is found by

$$\hat{\mathcal{T}} = A^{\dagger} U_s$$

where $A^{\dagger}$ is the Moore-Penrose pseudoinverse of A.

- By substituting $\hat{\mathcal{T}}$ to the above optimization problem we find an estimate for θ

$$\hat{\theta} = \arg\min_{\theta} Tr\{\Pi_A^{\perp} \hat{U}_s \hat{\Lambda}_s \hat{U}_s^H\},$$

where $\Pi_A^{\perp}$ is a projection matrix $I - \Pi_A = A[A^H A]^{-1} A^H$.

- We may weight the eigenvectors and obtain better estimates by

$$\hat{\theta} = \arg\min_{\theta} Tr\{\Pi_A^{\perp} \hat{U}_s W \hat{U}_s^H\}$$

Weighted Subspace Fitting (WSF) method

- This is equivalent to minimizing

$$[\hat{\theta}, \hat{\mathcal{T}}] = \arg\min_{\theta, \mathcal{T}} ||U_s W^{1/2} - A\mathcal{T}||_F^2$$

- Weighting matrix should be such that the variance of the estimation error is minimized.
- The optimum weighting matrix is

$$W_{opt} = (\Lambda_s - \sigma^2 I)^2 \Lambda_s^{-1},$$

where Λ_s are the eigenvalues of Σ correponding to U_s.

Weighted Subspace Fitting (WSF) method

- However, we have to use

$$W_{opt} = (\hat{\Lambda}_s - \hat{\sigma}^2 I)^2 \hat{\Lambda}_s^{-1}$$

  because of the unknowns.

- Estimate of the noise variance may be found via arithmetic mean of the $M - K'$ smallest eigenvalues of Σ.
- The method has lower computational complexity than Stochastic ML method but is asymptotically efficient (reaches CRLB).

Maximum Likelihood method in DoA Estimation

- Two different techniques depending on signal model
- Stochastic Maximum Likelihood (SML) assumes the signal and noise be random and gaussian.
- Deterministic Maximum Likelihood (DML) assumes that the signal is deterministic and noise are Gaussian.
- These methods can deal with coherent (highly correlated signals) resulting for example from multipath or jamming.
- The error criteria to be minimized have many local minima
- hence ML methods require a multidimensional search which has a high computational cost.
- EM algorithm and improved versions of it such as SAGE are often used to find the optimum.
- Other techniques include Iterative quadratic ML (IQML) method and Newton-type techniques.

Deterministic Maximum Likelihood (DML) method in DoA Estimation

- Signals are unknown deterministic waveforms and the noise is a zero mean spatially and temporally white circular Gaussian random process, i.e. $E[\boldsymbol{v}(n)\boldsymbol{v}^T(n)] = 0$ and $E[\boldsymbol{v}(n)\boldsymbol{v}^H(n)] = \sigma^2 I$.
- The observations $\mathbf{x}$ are distributed as $\mathcal{CN}(A\mathrm{s}, \sigma^2 I)$
- The pdf of one observation vector $\boldsymbol{x}(k)$ is

$$l(k) = \frac{1}{(\pi\sigma^2)^M} exp(-||\boldsymbol{x}(k) - A\boldsymbol{s}(k)||^2/\sigma^2)$$

DML method in DoA Estimation

- Measurements are independent, hence the likelihood function is the product

$$\boldsymbol{l}(\theta, \boldsymbol{s}, \sigma^2) = \prod_{k=1}^{N} \frac{1}{(\pi\sigma^2)^M} exp(-||\boldsymbol{x}(k) - A\boldsymbol{s}(k)||^2/\sigma^2)$$

- By ignoring terms independent of parameters of interest, we get the log-likelihood function

$$L(\theta, \boldsymbol{s}, \sigma^2) = M\ log(\sigma^2) + \frac{1}{\sigma^2 N}\sum_{k=1}^{N} ||\boldsymbol{x}(k) - A\boldsymbol{s}(k)||^2$$

with respect to $\sigma^2, \boldsymbol{\theta}, \boldsymbol{s}(k)$

DML method in DoA Estimation

- The DML estimate for the DOAs is of the form
$$\hat{\theta} = \arg\{\min_{\theta} \mathrm{Tr}\{\Pi_A^{\perp}\hat{\Sigma}\}\}.$$
where $\hat{\Sigma} = \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{x}(k)\boldsymbol{x}(k)^H$
- Measurements are projected onto orthogonal (null subspace of A) subspace and power is minimized. Or one could also say that the projection to signal subspace is maximized.
- In addition to θ other parameters may be estimated simultaneously
- ML estimates for the signals and noise variance can be calculated by
$$\hat{\sigma}^2 = \frac{1}{M-K}\mathrm{Tr}\{\Pi_A^{\perp}\hat{\Sigma}\} \quad \text{and} \quad \hat{\mathbf{s}}(n) = A^{\dagger}\mathbf{x}(n).$$
where $A^{\dagger}$ is the Moore-Penrose pseudoinverse of A.

DML method in DoA Estimation

- Typically $\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i)$ but other estimates of covariance may also be used.
- The method is the same as the least squares fitting method (which is ML in Gaussian case)
- DML method does not achieve the Cramer-Rao lower bound when $N \to \infty$ for finite M. The number of unknowns $s(k)$ increases when the number of observations increases.
- Non-linear multidimensional optimization problem. One could use Gauss-Newton methods or alternating projection method such as SAGE (EM-algorithm).

Spectral methods (BF, MVDR, MUSIC) used as initial estimates to Gauss Newton method

Stochastic Maximum Likelihood method

- In SML the signal vector $\mathbf{s}$ is modeled as a zero mean and temporally white circular Gaussian random process distributed as

$$\mathcal{CN}(0, \Sigma_s)$$

where circular means that $E[\boldsymbol{s}(k)\boldsymbol{s}(k)^T] = 0$.

- The noise are assumed to be Gaussian distributed as well, hence the observations $\mathbf{x}$ are temporally white and distributed as

$$\mathcal{CN}(0, \Sigma)$$

- Its second order moments are given by

$$E\{\mathbf{s}(l)\mathbf{s}^H(k)\} = \Sigma_s \delta_{lk} \quad \text{and} \quad E\{\mathbf{s}(l)\mathbf{s}^T(k)\} = 0$$

where δ_{lk} is the Kronecker delta and Σ_s signal covariance matrix.

Stochastic Maximum Likelihood method

- The above expression implies that the real and imaginary parts of any marginal component of $\mathbf{s}(n)$ are independent and identically distributed.
- Some key assumptions
  - the rank of Σ_s is $K' \leq K$.
  - to ensure parameter identifiability $K < (M + K')/2$.
  - noise $\mathbf{v}(n)$ is assumed to be zero mean spatially and temporally white circular Gaussian process

$$E\{\mathbf{v}(l)\mathbf{v}^H(k)\} = \sigma^2 I \delta_{lk} \quad \text{and} \quad E\{\mathbf{v}(l)\mathbf{v}^T(k)\} = 0.$$

  - signal and noise are mutually independent.
- From Gaussianity it follows that that the negative log likelihood function of the snapshot data $\mathbf{x}(1), \ldots, \mathbf{x}(n)$ is proportional to

$$l(\theta, \Sigma_s, \sigma^2) = \text{Tr}[\Pi_A^\perp \hat{\Sigma}]$$

Stochastic Maximum Likelihood method

- We can write the likelihood:

$$l(\theta, \Sigma_s, \sigma^2) = \log |A\Sigma_s A^H + \sigma^2 I| + \text{Tr}\{(A\Sigma_s A^H + \sigma^2 I)^{-1}\hat{\Sigma}\}$$

  where $|\cdot|$ stands for the determinant, Tr denotes the trace, and $\theta = [\theta_1, \dots, \theta_K]^T$ and $\Sigma = A\Sigma_s A^H + \sigma^2 I$.

- For fixed θ, the minimum with respect to σ^2 and Σ_s can be shown to be

$$\hat{\sigma}^2(\theta) = \frac{1}{M-K}\text{Tr}\{\Pi_A^{\perp}\hat{\Sigma}\}$$

$$\hat{\Sigma}_s(\theta) = A^{\dagger}(\hat{\Sigma} - \hat{\sigma}^2(\theta)I)A^{\dagger H}.$$

- Finally, the estimate of θ is

$$\hat{\theta}_{SML} = \arg\{\min_{\theta} \log |A\hat{\Sigma}_s(\theta)A^H + \hat{\sigma}^2(\theta)I|\}.$$

Stochastic Maximum Likelihood method

- Converges for large N
- Asymptotically SML reaches the CRLB for Gaussian $s(k)$, it is consistent and asymptotically efficient. To conclude SML is optimal.
- For small number of sensors, low SNR and highly correlated signals SML estimates are significantly better that deterministic ML estimates (DML).

CRLB for DoA Estimation

- Let us have an array of M sensors and K signals with AoA's $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_K]$ impinging the array. Signals are assumed to be stochastic, hence we talk about stochastic CRLB. The number of snapshots is N.
- Cramer-Rao Lower Bound gives the lowest possible variance for an unbiased estimator. It is defined as the inverse of the Fisher Information Matrix.
  Slepian-Bangs formula can be used in Gaussian models

CRLB for DoA Estimation

- By employing asymptotical analysis of ML estimate of $\boldsymbol{\theta}$ and by using the fact that the covariance matrix of the ML estimator asymptotically coincides with the CRB, we find the CRB (as function of the incident angles $\boldsymbol{\theta}$) as

$$CRLB(\boldsymbol{\theta}) = \frac{\sigma_v^2}{2N}\left\{Re\left[(\boldsymbol{\Delta}^H \boldsymbol{\Pi}_v \boldsymbol{\Delta}) \odot (\boldsymbol{\Sigma}_s \mathbf{A}^H \boldsymbol{\Sigma}_x^{-1} \mathbf{A} \boldsymbol{\Sigma}_s)^T\right]\right\}^{-1} \quad (2)$$

where $\boldsymbol{\Delta} = [\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K]$ with $\boldsymbol{\delta}_k = \left.\frac{d\mathbf{a}(\theta)}{d\theta}\right|_{\theta=\theta_k}$, is a $M \times K$ vector containing the derivative of the array steering vector with respect to the angle, evaluated at a certain DoA θ_k, $\mathbf{A}$ is the $M \times K$ array steering matrix, $\boldsymbol{\Sigma}_s$ is the $K \times K$ signal covariance matrix, $\boldsymbol{\Sigma}_x$ is the $M \times M$ array covariance matrix, N is the number of snapshots, σ_v^2 is the AWG noise power and $\odot$ denotes the Hadamard-Schur product, i.e., element-wise multiplication of matrices.

CRLB for DoA Estimation

- $Re[\bullet]$ means that real part is taken
- Moreover, $\mathbf{\Pi}_v = \mathbf{I} - \mathbf{\Pi}_s$ is the $M \times M$ projection matrix to noise subspace where $\mathbf{\Pi}_s = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ is the $M \times M$ projection matrix to signal subspace.
- The following approximate formula was found by Stoica and Nehorai:

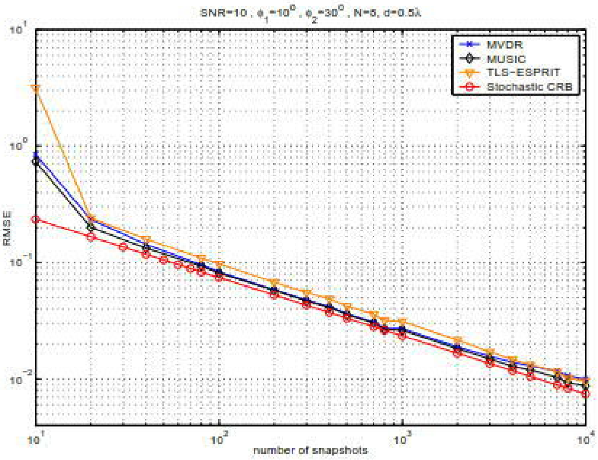
$$CRLB(\infty, \infty) = \frac{6}{M^3 N} \begin{bmatrix} \frac{1}{SNR_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{SNR_K} \end{bmatrix}$$

Approximation for Large arrays and high SNR, for ULAs only

for sufficiently large array sizes M and number of snapshots N. SNR_i denotes the SNR for the ith signal.

EX: CRLB for DoA Estimation

- CRLB as function of snapshots N for 8-element ULA. $SNR = 20$ dB, 2 sources at angles 10 and 30 degrees. Error variances of MVDR, MUSIC and ESPRIT are shown as well.



Model Order / Number of Signals Estimation

- Different ways to estimate the number of signals:
  - Hypothesis testing: Generalized Likelihood Ratio Test (GLRT)
  - Akaike Information Criterion (AIC)
  - Minimum Description Length (MDL) by Rissanen
- Typically eigenvalues of the covariance matrix are employed
- MDL technique presented here is consistent
- MDL estimates the number of signals by finding an integer $k \in \{0, 1, \ldots, M-1\}$ which minimizes the criterion

$$MDL(k) = -\log \left(\frac{\left(\prod_{i=k+1}^{M} \hat{\lambda}_i\right)^{1/(M-k)}}{\frac{1}{M-k}\sum_{i=k+1}^{M} \hat{\lambda}_i} \right)^{(M-k)N} + \frac{1}{2}k(2M-k)\log N,$$

where $\hat{\lambda}_i, i = 1, \ldots, M$ are the eigenvalues of the sample covariance matrix.

Model Order / Number of Signals Estimation

- Signals and noise are assumed to be Gaussian in the derivation of the criterion.
- The ratio of geometric and arithmetic mean of the eigenvalues tests the equality of the noise subspace eigenvalues (symmetry of the noise subspace)
- Noise power may be estimated as

$$\hat{\sigma}_v^2 = \frac{1}{M-K}\sum_{i=K+1}^{M} \hat{\lambda}_i,$$

i.e, the arithmetic mean of the noise subspace eigenvalues and the noise covariance matris estimate is $\hat{\Sigma}_v = \sigma_v^2 I$

In some special cases Bayesian Information Criterion (BIC) has the same form as MDL. This array processing model is one such case.

Coherent or correlated signals

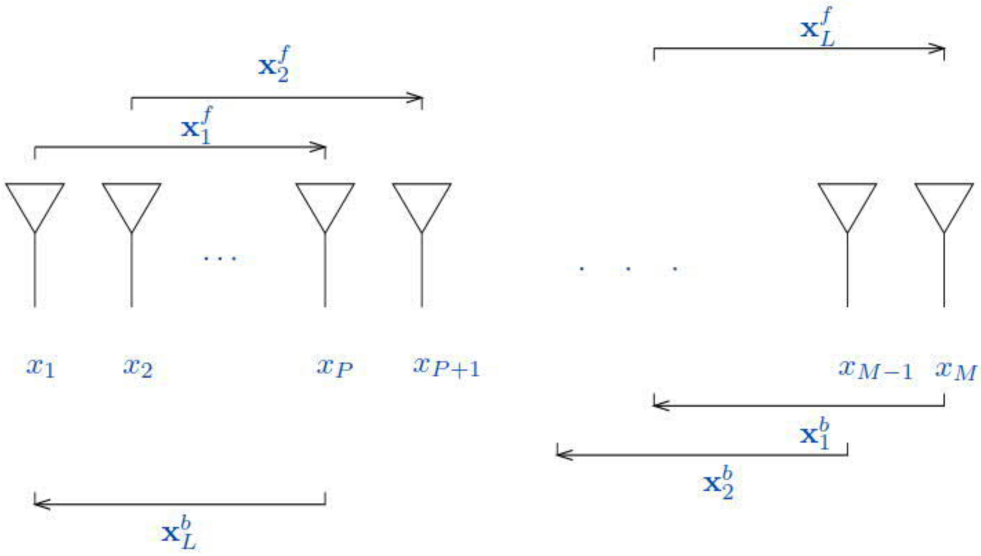
- Multipath propagation or intentional jamming may cause the received signal from different θ_i to be perfectly correlated
- Consequently, signal covariance matrix may not be of full rank
- Maximum Likelihood methods can deal with such situations
- Subspace methods obviously perform poorly
- Weighted subspace fitting can deal with coherent signals as well

Spatial Smoothing

- In the case of ULA, the DOAs can, however, be estimated by using spatial smoothing preprocessing step.
- A preprocessing step called spatial smoothing is applied to build the rank of the array covariance matrix
  - Subdivide the array of M sensors into subarrays of P sensors
  - compute sample covariance matrices for the subarrays
  - Average the subarrays (forward-backward averaging)
  - Plug in the averaged covariance matrix resulting from spatial smoothing to your favourite subspace methdod
  - The penalty we pay is that only $2M/3$ sources may be resolved.
- Let $\mathbf{x}_l^f$ denote the received signals at the lth forward subarray i.e.

$$\mathbf{x}_l^f = (x_l, \ldots, x_{l+P-1})^T.$$

Spatial Smoothing



The forward/backward spatial smoothing scheme.

Spatial Smoothing

- Let $\mathbf{x}_l^b$ denote the complex conjugate of the received signals at the lth backward subarray

$$\mathbf{x}_l^b = (x_{M-l+1}^*, \ldots, x_{M-P-l+2}^*)^T.$$

- Let D^l denote the lth power of the diagonal matrix

$$D = \text{diag}\{e^{-j2\pi(d/\lambda)\cos(\theta_1)}, \ldots, e^{-j2\pi(d/\lambda)\cos(\theta_K)}\}.$$

- Adapting the same notation as before, we can model $\mathbf{x}_l^f$ and $\mathbf{x}_l^b$ as:

$$\mathbf{x}_l^f = AD^{(l-1)}\mathbf{s} + \mathbf{v}_l^f \quad \text{and} \quad \mathbf{x}_l^b = AD^{(l-1)}\left(D^{(M-1)}\mathbf{s}\right)^* + \mathbf{v}_l^b$$

where $A = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ with $\mathbf{a}(\theta_k)$ being the $P \times 1$ ($P > K$) array steering vector corresponding to the DOA of the kth signal component, and $\mathbf{v}_l^f$ and $\mathbf{v}_l^b$ are noise vectors.

Spatial Smoothing

- The forward-averaged spatially-smoothed covariance matrix Σ^f is defined as the sample mean of the subarray covariance matrices

$$\Sigma^f = \frac{1}{L}\sum_{l=1}^{L}\Sigma_l^f,$$

where $\Sigma_l^f = E\{\mathbf{x}_l^f \mathbf{x}_l^{f\,H}\}$ and $L = M - P + 1$.

- The backward averaged spatially smoothed covariance matrix is

$$\Sigma^b = \frac{1}{L}\sum_{l=1}^{L}\Sigma_l^b$$

where $\Sigma_l^f = E\{\mathbf{x}_l^b \mathbf{x}_l^{b\,H}\}$.

Spatial Smoothing

- The forward/backward spatial smoothed covariance matrix $\bar{\Sigma}$ is defined as

$$\bar{\Sigma} = \frac{\Sigma^f + \Sigma^b}{2}. \quad (3)$$

- it is possible to choose P such that the $P - K$ smallest eigenvalues of $\bar{\Sigma}$ are equal and the corresponding eigenvectors are orthogonal to the columns of the matrix A
- The DOAs of the coherent signals can be estimated using any subspace algorithm and an estimate of $\bar{\Sigma}$.

Thank you for your attention!

Aalto University
School of Engineering