E401: Advanced Communication Theory

Professor A. Manikas
Chair of Communications and Array Processing

Imperial College London

Multi-Antenna Wireless Communications

Array Receivers for SIMO and MIMO

Table of Contents

General Objective

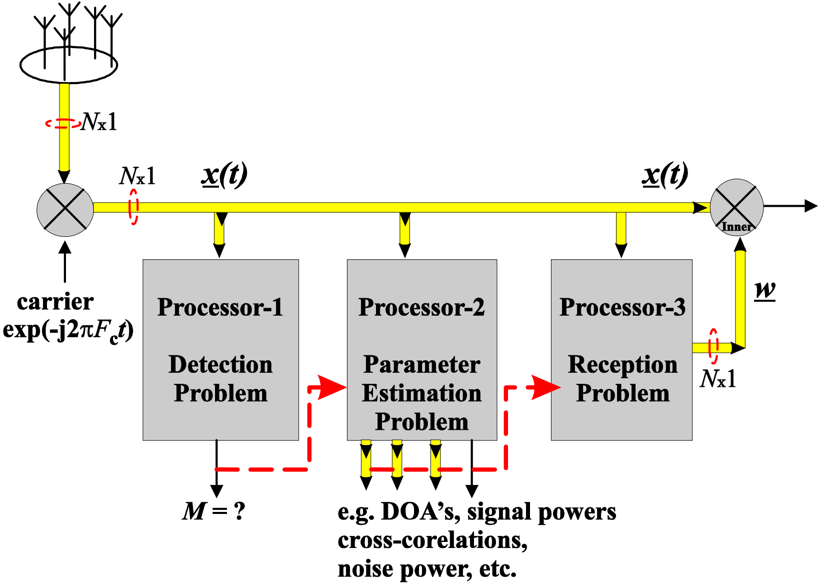
- By observing a vector-signal $\underline{x}(t) = \mathbb{S}\underline{m}(t)+\underline{n}(t)$ using an array system the aim is to obtain information about a signal environment.
- There are three general problems to solve.

1. The <u>Detection</u> problem: $M =?$
   (i.e. to detect the presence of M **co-channel** emitting sources)

2. The <u>Estimation</u> problem:
   to estimate various signal and channel parameters
   e.g. DOAs$=?\forall i$; $P_{m_i} = \mathcal{E}\{m_i^2(t)\} =?\forall i$; $P_{\mathrm{n}} = \sigma_{\mathrm{n}}^2 =?$;
   $\rho_{ij} = \mathcal{E}\{m_i(t)m_j^*(t)\} =?\forall i, j$, with $i \neq j$
   polarization parameters, fading coefficients, signal spread.

3. The <u>Reception</u> problem:
   to receive one signal (desired signal) and suppress the remaining $M-1$ as unwanted cochannel interference

- These problems are highlighted in the following block structure:



General Problem Formulation M<N

- Consiner an observed $(N \times 1)$ complex signal-vector $\underline{x}(t)$ that is modelled as follows

$$\underline{x}(t) \triangleq \overbrace{\mathbb{S}(\underline{p})}^{N\times M} \cdot \overbrace{\underline{m}(t)}^{M\times 1} + \overbrace{\underline{\mathrm{n}}(t)}^{N\times 1} \tag{1}$$

Note that by observing $\underline{x}(t)$, its 2nd order statistics become known, i.e. the covariance matrix $\mathbb{R}_{xx}$ is known, where

$$\mathbb{R}_{xx} = \mathcal{E}\{\underline{x}(t).\underline{x}(t)^H\} \tag{2}$$

- Estimate $M, p_1, p_2, ..., p_M, \mathbb{R}_{mm}, \sigma_{\mathrm{n}}^2$, etc.

$$\text{where} \begin{cases} \mathbb{S} \triangleq \mathbb{S}(\underline{p}) = \left[\underline{S}(p_1), \quad \underline{S}(p_2), \quad ..., \quad \underline{S}(p_M)\right] \text{ - (unknown)} \\ \underline{m}(t): \quad \text{message signal-vector - (unknown)} \\ \mathbb{R}_{mm}: \quad \text{2nd order statistics of } \underline{m}(t) \text{ - (unknown)} \\ \underline{n}(t): \quad \text{AWGN vector - (power } \sigma_n^2 \text{ unknown)} \end{cases}$$

$$\text{with} \begin{cases} \underline{p} = \quad \text{the vector of generic (unknown) parameters } p_1, p_2, ..., p_M \\ N = \quad \textbf{known} \text{ (this is a system parameter)} \\ M = \quad \textbf{unknown} \text{ (this is a channel parameter - number of signals)} \\ \qquad \text{with } M < N \text{ (later this condition will be removed)} \end{cases}$$

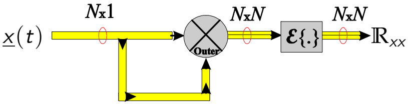
Array Covariance Matrix

Theoretical Covariance Matrix

- If the $(N \times 1)$ vector-signal $\underline{x}(t) = \mathbb{S}\underline{m}(t) + \underline{n}(t)$ is observed over infinite observation interval then its 2nd order statistics can be calculated. These are given by the theoretical covariance matrix $\mathbb{R}_{xx}$ which is an $(N \times N)$ complex matrix - always Hermitian. That is,

$$
\begin{aligned}
\mathbb{R}_{xx} &\triangleq \mathcal{E}\left\{\underline{x}(t)\underline{x}(t)^H\right\} \qquad (3)\\
&= \begin{bmatrix}
\mathcal{E}\{x_1(t)x_1(t)^*\}, & \mathcal{E}\{x_1(t)x_2(t)^*\}, & \ldots, & \mathcal{E}\{x_1(t)x_N(t)^*\} \\
\mathcal{E}\{x_2(t)x_1(t)^*\}, & \mathcal{E}\{x_2(t)x_2(t)^*\}, & \ldots, & \mathcal{E}\{x_2(t)x_N(t)^*\} \\
\ldots, & \ldots, & \ldots, & \ldots \\
\mathcal{E}\{x_N(t)x_1(t)^*\}, & \mathcal{E}\{x_N(t)x_2(t)^*\}, & \ldots, & \mathcal{E}\{x_N(t)x_N(t)^*\}
\end{bmatrix} \\
&= \mathcal{E}\left\{(\mathbb{S}\underline{m}(t) + \underline{n}(t)) . (\mathbb{S}\underline{m}(t) + \underline{n}(t))^H\right\} \\
&= \mathcal{E}\left\{\mathbb{S}\underline{m}(t)\underline{m}(t)^H\mathbb{S}^H + \underline{n}(t)\underline{n}(t)^H + \mathbb{S}\underline{m}(t)\underline{n}(t)^H + \underline{n}(t)\underline{m}(t)^H\mathbb{S}^H\right\} \\
&= \mathbb{S}.\underbrace{\mathcal{E}\left\{\underline{m}(t)\underline{m}(t)^H\right\}}_{\triangleq\mathbb{R}_{mm}}.\mathbb{S}^H + \underbrace{\mathcal{E}\left\{\underline{n}(t)\underline{n}(t)^H\right\}}_{\triangleq\mathbb{R}_{nn}} + \mathbb{S}\underbrace{\mathcal{E}\left\{\underline{m}(t)\underline{n}(t)^H\right\}}_{=\mathbb{O}_{M,N}} + \underbrace{\mathcal{E}\left\{\underline{n}(t)\underline{m}(t)^H\right\}}\mathbb{S}^H \\
&= \mathbb{S}\,.\mathbb{R}_{mm}.\,\mathbb{S}^H + \mathbb{R}_{nn} \qquad (4)
\end{aligned}
$$

- i.e.



$$\mathbb{R}_{xx} \triangleq \mathcal{E}\left\{\underline{x}(t)\underline{x}(t)^H\right\}$$

$$= \mathbb{S}.\mathbb{R}_{mm}.\mathbb{S}^H + \mathbb{R}_{\text{nn}}$$

where

$$\mathbb{R}_{mm} \triangleq \mathcal{E}\left\{\underline{m}(t).\underline{m}(t)^H\right\} = \text{2nd order statistics of } \underline{m}(t) \text{ (unknown)}$$

$$= \begin{bmatrix} \underbrace{\mathcal{E}\{m_1(t).m_1(t)^*\}}_{\mathcal{E}\{m_1(t)^2\}=P_1}, & \mathcal{E}\{m_1(t).m_2(t)^*\}, & ..., & \mathcal{E}\{m_1(t).m_M(t)^*\} \\ \mathcal{E}\{m_2(t).m_1(t)^*\}, & \underbrace{\mathcal{E}\{m_2(t).m_2(t)^*\}}_{\mathcal{E}\{m_2(t)^2\}=P_2}, & ..., & \mathcal{E}\{m_2(t).m_M(t)^*\} \\ \vdots & \vdots & \vdots, & \vdots \\ \mathcal{E}\{m_M(t).m_1(t)^*\}, & \mathcal{E}\{m_M(t).m_2(t)^*\}, & ..., & \underbrace{\mathcal{E}\{m_M(t).m_M(t)^*\}}_{\mathcal{E}\{m_M(t)^2\}=P_M} \end{bmatrix}$$

$$= \text{an } (M \times M) \text{ complex matrix (always Hermitian) - unknown}$$

$$\begin{aligned}
\mathbb{R}_{\mathrm{nn}} &\triangleq \mathcal{E}\left\{\underline{\mathrm{n}}(t).\underline{\mathrm{n}}(t)^H\right\} \text{ is 2nd order statistics of } \underline{\mathrm{n}}(t) \\
&= \begin{bmatrix}
\underbrace{\mathcal{E}\left\{\mathrm{n}_1(t).\mathrm{n}_1(t)^*\right\}}_{\mathcal{E}\{\mathrm{n}_1(t)^2\}=P_{\mathrm{n}_1}}, & \underbrace{\mathcal{E}\left\{\mathrm{n}_1(t).\mathrm{n}_2(t)^*\right\}}_{0}, & \ldots, & \underbrace{\mathcal{E}\left\{\mathrm{n}_1(t).\mathrm{n}_N(t)^*\right\}}_{0} \\
\underbrace{\mathcal{E}\left\{\mathrm{n}_2(t).\mathrm{n}_1(t)^*\right\}}_{0}, & \underbrace{\mathcal{E}\left\{\mathrm{n}_2(t).\mathrm{n}_2(t)^*\right\}}_{\mathcal{E}\{\mathrm{n}_2(t)^2\}=P_{\mathrm{n}_2}}, & \ldots, & \underbrace{\mathcal{E}\left\{\mathrm{n}_2(t).\mathrm{n}_N(t)^*\right\}}_{0} \\
\ldots, & \ldots, & \ldots, & \ldots \\
\underbrace{\mathcal{E}\left\{\mathrm{n}_N(t).\mathrm{n}_1(t)^*\right\}}_{0}, & \underbrace{\mathcal{E}\left\{\mathrm{n}_N(t).\mathrm{n}_2(t)^*\right\}}_{0}, & \ldots, & \underbrace{\mathcal{E}\left\{\mathrm{n}_N(t).\mathrm{n}_N(t)^*\right\}}_{\mathcal{E}\{\mathrm{n}_N(t)^2\}=P_{\mathrm{n}_N}}
\end{bmatrix} \\
&= \sigma_{\mathrm{n}}^2 \mathbb{I}_N \qquad (5) \\
&= \text{an } (N \times N) \text{ complex matrix (always Hermitian) - unknown}
\end{aligned}$$

- Note that, because we have assumed isotropic AWGN noise,

$$P_{\mathrm{n}_1} = P_{\mathrm{n}_2} = \ldots = P_{\mathrm{n}_N} = \sigma_{\mathrm{n}}^2 \qquad (6)$$

Practical Covariance Matrix

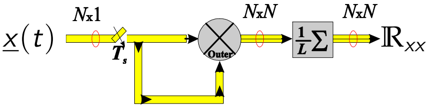
- Consider that the signal $\underline{x}(t) = \mathbb{S}\underline{m}(t) + \underline{\mathrm{n}}(t)$ is observed over finite observation interval equivalent to L snapshots.
- These L observations (snapshots) at times t_1, t_2, ..., t_L (i.e. finite observation interval) are denoted as $[\underline{x}(t_1), \underline{x}(t_2), ..., \underline{x}(t_L)]$ and represented by the $N \times L$ complex matrix $\mathbb{X}$ i.e.

$$\begin{aligned} \mathbb{X} &\triangleq [\underline{x}(t_1), \underline{x}(t_2), ..., \underline{x}(t_L)] && (7a)\\ &= [\mathbb{S}.\underline{m}(t_1) + \underline{\mathrm{n}}(t_1), \quad \mathbb{S}.\underline{m}(t_2) + \underline{\mathrm{n}}(t_2), \quad ..., \quad \mathbb{S}.\underline{m}(t_L) + \underline{\mathrm{n}}(t_L)] \\ &= \mathbb{S}.\mathbb{M} + \mathbb{N} && (7b) \end{aligned}$$

$$with \begin{cases} \mathbb{S} = [\underline{S}_1, \underline{S}_2, ..., \underline{S}_M] & (N \times M) \\ \mathbb{M} = [\underline{m}(t_1), \underline{m}(t_2), ..., \underline{m}(t_L)] & (M \times L) \\ \mathbb{N} = [\underline{\mathrm{n}}(t_1), \underline{\mathrm{n}}(t_2), ..., \underline{\mathrm{n}}(t_L)] & (N \times L) \end{cases} \quad (8)$$

where the matrices $\mathbb{S}$, $\mathbb{M}$ and $\mathbb{N}$ (as well as the dimension M) are unknown

- In this case the 2nd order statistics of $\underline{x}(t)$ are estimated by the practical covariance matrix $\mathbb{R}_{xx}$
- Practical Model:



$$\mathbb{R}_{xx} = \frac{1}{L}\sum_{l=1}^{L} \underline{x}(t_l)\underline{x}(t_l)^H \quad (9)$$

$$\text{i.e. } \mathbb{R}_{xx} = \frac{1}{L}\mathbb{X}\mathbb{X}^H = \mathbb{S}\underbrace{\frac{1}{L}\mathbb{M}\mathbb{M}^H}_{=\mathbb{R}_{mm}}\mathbb{S}^H + \underbrace{\frac{1}{L}\mathbb{N}\mathbb{N}^H}_{=\mathbb{R}_{nn}}$$

$$= \mathbb{S}\,.\mathbb{R}_{mm}.\,\mathbb{S}^H + \mathbb{R}_{nn} \quad (10)$$

N.B.:

- In an array system the matrix $\mathbb{R}_{xx}$ (theoretical or practical) contains all the **geometrical** **and** **other information** about the various sources relative to the array.
- Remember that, sometimes, we will use $\widehat{\mathbb{R}}_{xx}$ to denote the practical/estimated covariance matrix of the vector signal $\underline{x}(t)$.

Generating L Snapshots having a given Covariance Matrix

- To generate L snapshots of $\underline{x}(t)$ having a predefined covariance matrix $\mathbb{R}_{xx}$ the vectors $\underline{x}(t_l)$ for $l = 1, 2, .., L$ should be generated using the following expression

$$\boxed{\underline{x}(t_l) = \mathbb{E}\,\mathbb{D}^{\frac{1}{2}}\,\underline{z}(t_l)} \tag{11}$$

where

  - $\mathbb{E}$ and $\mathbb{D}$ are the eigenvector-matrix and eigenvalue-matrix of $\mathbb{R}_{xx}$, and
  - $\underline{z}(t_l) \in \mathcal{C}^N$ is a Gaussian random complex vector of N elements of zero mean and variance 1, i.e.

$$\mathcal{E}\left\{\underline{z}(t_l).\underline{z}(t_l)^H\right\} = \mathbb{I}_N \tag{12}$$

- That is,

$$\begin{aligned} \mathbb{X} &= \left[\underline{x}(t_1), \quad \underline{x}(t_2), \quad ..., \quad \underline{x}(t_L)\right] \\ &= \left[\mathbb{E}\,\mathbb{D}^{\frac{1}{2}}\underline{z}(t_1), \quad \mathbb{E}\,\mathbb{D}^{\frac{1}{2}}\underline{z}(t_2), \quad ..., \quad \mathbb{E}\,\mathbb{D}^{\frac{1}{2}}\underline{z}(t_L)\right] \end{aligned} \tag{13}$$

Proof.

- Let $\underline{z} \in \mathcal{C}^N$ such as $\mathcal{E}\{\underline{z}.\underline{z}^H\} = \mathbb{I}_N$. Then

$$\mathbb{R}_{xx} = \mathcal{E}\{\underline{x}.\underline{x}^H\} \tag{14}$$

- However,

$$\begin{aligned} \mathbb{R}_{xx} &= \mathbb{E}\mathbb{D}\mathbb{E}^H = \mathbb{E}\mathbb{D}^{\frac{1}{2}}\mathbb{D}^{\frac{1}{2}}\mathbb{E}^H = \mathbb{E}\mathbb{D}^{\frac{1}{2}}\mathbb{I}_N\mathbb{D}^{\frac{1}{2}}\mathbb{E}^H \\ &= \mathbb{E}\mathbb{D}^{\frac{1}{2}}\underbrace{\mathcal{E}\{\underline{z}.\underline{z}^H\}}_{=\mathbb{I}_N}\mathbb{D}^{\frac{1}{2}}\mathbb{E}^H \\ &= \mathcal{E}\{\underbrace{\mathbb{E}\mathbb{D}^{\frac{1}{2}}\underline{z}}_{\underline{x}}.\underbrace{\underline{z}^H\mathbb{D}^{\frac{1}{2}}\mathbb{E}^H}_{\underline{x}^H}\} \end{aligned} \tag{15}$$

- By comparing Equation 15 and 14 we have

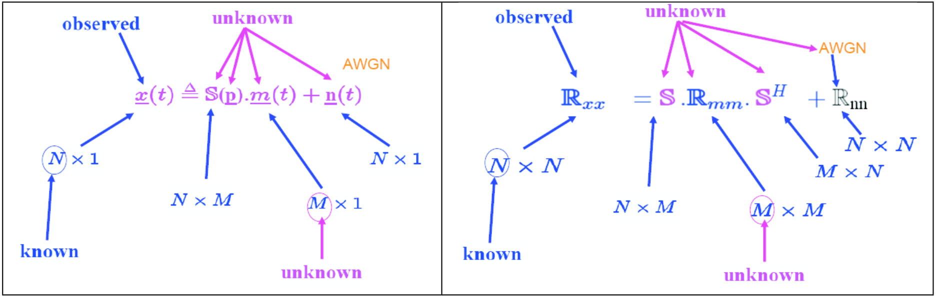
$$\boxed{\underline{x} = \mathbb{E}\,\mathbb{D}^{\frac{1}{2}}\,\underline{z}}$$

□

N.B.:

- In a similar fashion we can generate L snapshots of $\underline{m}(t)$, $\underline{n}(t)$ or any other vector-signal with known cov. matrix.

Summary - General Problem Formulation



- Condition: $M < N$
- Estimate $M, \underline{p} = [p_1, p_2, ..., p_M]^T, \mathbb{R}_{mm}, \sigma_n^2$, etc where p_i is a parameter of interest associated with the i^{th} source.

The ‘Detection’ Problem: Basic Detection Theory

Problem

- *To determine the parameter M*

  i.e. to determine the number of signals and thus the dimensions of the vectors/matrices $\underline{m}(t)$, $\mathbb{S}$, $\mathbb{R}_{mm}$ and $\mathbb{M}$

- *In other words, to detect how many emitting sources/transmitters are present in an array environment*

  i.e. to detect the presence of M sources

Solution

- *To use optimum Detection and Decision Theory*
- *Note-1: The main optimum decision criteria are summarised in Appendix-A (at the end of this handout).*
- *Note-2: The ML decision criterion will be employed in this section.*

Hypothesis Testing

Definition (Hypothesis)

A Hypothesis $\triangleq$ a statement of a possible condition

Definition (Hypethesis Testing)

To choose one from a number (two or more) hypotheses

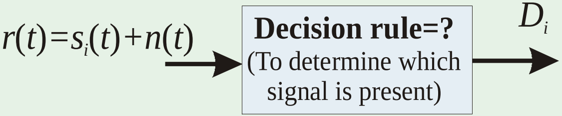
Example (1: Detection of a signal $s(t)$ in the presence of noise)

- we have an observed signal $r(t)$
- we define two hypotheses H_1 and H_2

$$\text{Hypethesis Testing:} \begin{cases} H_1: \quad r(t) = s(t) + n(t) & \text{i.e. the signal is present} \\ \\ H_2: \quad r(t) = n(t) & \text{i.e.the signal is not present} \end{cases}$$

Example (2: Hypothesis Testing in an M-ary Comm System)

- In an M-ary Comm. System we have M hypotheses
- The aim is to design a receiver which operates on an observed signal $r(t)$ and chooses one of the following M hypotheses:



$$\text{Hypothesis Testing}\begin{cases} H_1: & r(t)=s_1(t)+n(t), \\ H_2: & r(t)=s_2(t)+n(t), \\ \ldots & \ldots \\ H_M: & r(t)=s_M(t)+n(t), \end{cases} \tag{16}$$

where $s_i(t)=$ one of M signals (channel symbols)

Terminology

Consider an observed signal $r(t)$ and M hypotheses H_1, H_2,..., H_M.

- **A priori probabilities** $\triangleq$

$$\Pr(H_1), \Pr(H_2), ..., \Pr(H_M)$$

(these are calculated **BEFORE** the experiment is performed)

- **A posterior probabilities** $\triangleq$

$$\Pr(H_1/r), \Pr(H_2/r), ..., \Pr(H_M/r)$$

That is, if r =observation variable
then we have M Conditional Probabilities

$$\Pr(H_i/r), \forall i \in [1, \ldots, M]$$

known as a POSTERIOR PROBABILITIES
(since these are calculated **AFTER** the experiment is performed).

- $\Pr(H_i/r)$ $\forall i$: difficult to find. A more natural approach is to find

$$\Pr(r/H_i), \forall i \tag{17}$$

since in general $\text{pdf}_{r/H_i}, \forall i$

- are known or
- can be found

Definition (Likelihood Functions (LF))

the M conditional probability density functions $\text{pdf}_{r/H_i}(r), \forall i$, i.e.

$$\text{pdf}_{r/H_1}(r), \text{pdf}_{r/H_2}(r), ..., \text{pdf}_{r/H_M}(r) \tag{18}$$

are known as "Likelihood Functions"

Definition (Likelihood Ratio (LR))

The ratio

$$\frac{\text{pdf}_{r/H_i}(r)}{\text{pdf}_{r/H_j}(r)} \text{ for } i \neq j \tag{19}$$

is known as "Likelihood Ratio"

Hypothesis Testing re-Defined

- Note that the statistics of the observed signal $r(t)$ are different for different hypotheses.

Definition (Hypothesis Testing re-defined)

If the distributions of the observed signal $r(t)$ for various hypotheses are **known**

then the problem of choosing one of many (say M) hypotheses is **translated** to make a decision about one of the M distributions after having observed $r(t)$.

This is called 'Hypothesis Testing'

Antenna Array: Hypothesis Testing

- Based on the model of the signal received by an antenna array system (see Equ 1) we can define the following set of hypotheses:

$$\left\{\begin{array}{lll} H_1: & \underline{x}(t) = n(t); & \mathrm{LF}^{(0)} = \mathrm{pdf}_{\underline{x}(t)|H_1} \\ H_2: & \underline{x}(t) = \underline{S}_1 m_1(t) + n(t); & \mathrm{LF}^{(1)} = \mathrm{pdf}_{\underline{x}(t)|H_2} \\ H_3: & \underline{x}(t) = \sum_{i=1}^{2} \underline{S}_i m_i(t) + n(t); & \mathrm{LF}^{(2)} = \mathrm{pdf}_{\underline{x}(t)|H_3} \\ \cdots & \cdots & \cdots \\ H_N: & x(t) = \sum_{i=1}^{N-1} \underline{S}_i m_i(t) + n(t); & \mathrm{LF}^{(N-1)} = \mathrm{pdf}_{\underline{x}(t)|H_N} \end{array}\right\} \tag{20}$$

- Then we may use the ML decision rule, i.e.

$$\text{Maximum Likehood (ML)} = \max_k \left\{\mathrm{LF}^{(k)}\right\}$$

- Equation 20, can be rewritten in an equivalent way as follows:

$$\left\{\begin{array}{llll} H_1: & \mathbb{R}_{xx}^{(0)} = \mathbb{R}_{nn}; & \mathrm{LF}^{(0)} = \mathrm{pdf}_{\underline{x}(t)|H_1} \\ H_2: & \mathbb{R}_{xx}^{(1)} = \underbrace{P_1 \underline{S}_1 \underline{S}_1^H}_{\mathbb{R}_{signals}} + \mathbb{R}_{nn}; & \mathrm{LF}^{(1)} = \mathrm{pdf}_{\underline{x}(t)|H_2} \\ H_3: & \mathbb{R}_{xx}^{(2)} = \underbrace{\sum_{i=1}^{2} P_i \underline{S}_i \underline{S}_i^H}_{\mathbb{R}_{signals}} + \mathbb{R}_{nn}; & \mathrm{LF}^{(2)} = \mathrm{pdf}_{\underline{x}(t)|H_3} \\ \cdots & \cdots & \cdots \\ H_N: & \mathbb{R}_{xx}^{(N-1)} = \underbrace{\sum_{i=1}^{N-1} P_i \underline{S}_i \underline{S}_i^H}_{\mathbb{R}_{signals}} + \mathbb{R}_{nn}; & \mathrm{LF}^{(N-1)} = \mathrm{pdf}_{\underline{x}(t)|H_N} \end{array}\right\} \quad (21)$$

N.B.:

- The $\mathrm{LF}^{(k)}$ for $k = 0, 2, ..., N-1$ is a function of the eigenvalues of the matrices $\mathbb{R}_{signals}$ and, consequently, of the eigenvalues of $\mathbb{R}_{xx}^{(k)}$
- $\mathrm{ML} = \max_k \left\{ \mathrm{LF}^{(k)} \right\}$

Infinite Observation Interval (Infinity snapshots)

- Based on $\underline{x}(t)$ we can form the matrix $\mathbb{R}_{xx}$ (representing the statistics of $\underline{x}(t)$):

$$\mathbb{R}_{xx} \triangleq \mathcal{E}\left\{\underline{x}(t)\underline{x}(t)^H\right\} = \underbrace{\mathbb{S} \cdot \mathbb{R}_{mm} \cdot \mathbb{S}^H}_{=\mathbb{R}_{\text{signals}}} + \underbrace{\mathbb{R}_{\text{nn}}}_{=\sigma_{\text{n}}^2 \mathbb{I}_N} \quad (22)$$

  When the number of sources M is smaller than the number of system dimensions N (e.g. number of array-sensors) then the determinant of the $\mathbb{R}_{\text{signals}}$ is equal to zero
  i.e.

$$\text{if } M < N \Rightarrow \det(\mathbb{R}_{\text{signals}}) = 0 \quad (23)$$

  This is due to the fact that the presence of an emitting source increases the rank of the matrix $\mathbb{R}_{\text{signals}}$ by one.i.e.

$$\begin{aligned} & rank\left\{\mathbb{R}_{\text{signals}}\right\} = M && (24) \\ \Rightarrow \quad & rank\left\{\mathbb{R}_{xx} - \sigma_{\text{n}}^2 \mathbb{I}_N\right\} = M && (25) \end{aligned}$$

- However, using eigen-decomposition of $\mathbb{R}_{\mathbf{xx}}$ we have

$$\mathbb{R}_{\mathbf{xx}} = \mathbb{E}.\mathbb{D}.\mathbb{E}^H \tag{26}$$

where

$$\begin{aligned} \mathbb{D} &= \begin{bmatrix} d_{\mathbf{1}}, & 0, & ..., & 0, & 0, & 0, & ..., & 0 \\ 0, & d_{\mathbf{2}}, & ..., & 0, & 0, & 0, & ..., & 0 \\ ..., & ..., & ..., & ..., & ..., & ..., & ..., & ... \\ 0, & 0, & ..., & d_M, & 0, & 0, & ..., & 0 \\ 0, & 0, & ..., & 0, & d_{M+1}, & 0, & ..., & 0 \\ 0, & 0, & ..., & 0, & 0, & d_{M+2}, & ..., & 0 \\ ..., & ..., & ..., & ..., & ..., & ..., & ..., & ... \\ 0, & 0, & ..., & 0, & 0, & 0, & ..., & d_N \end{bmatrix} \\ &= (N \times N) \text{ matrix} \end{aligned} \tag{27}$$

with $d_\ell, \ell = 1, 2, ...N$, denoting the ℓ-th eigenvalue of $\mathbb{R}_{xx}$ and

$$d_{\mathbf{1}} > d_2 > ... > d_N > 0 \tag{28}$$

- An alternative way to express $\mathbb{R}_{\mathbf{xx}}$ as the addition of signals and noise covariance matrices and then to eigen-decompose the signal covariance matrix $\mathbb{R}_{signals}$. That is,

$$\begin{aligned}\mathbb{R}_{xx} &= \mathbb{R}_{signals} + \mathbb{R}_{noise} \\ &= \mathbb{E}.\underline{\underline{\Lambda}}.\mathbb{E}^H + \sigma_{\mathrm{n}}^2 \mathbb{I}_N = \mathbb{E}.\underline{\underline{\Lambda}}.\mathbb{E}^H + \sigma_{\mathrm{n}}^2 \underbrace{\mathbb{E}.\mathbb{E}^H}_{\mathbb{I}_N} \\ &= \mathbb{E}.\underbrace{\left(\underline{\underline{\Lambda}} + \sigma_{\mathrm{n}}^2 \mathbb{I}_N\right)}_{\mathbb{D}}.\mathbb{E}^H \end{aligned} \tag{29}$$

$$\text{where } \underline{\underline{\Lambda}} = \begin{bmatrix} \lambda_1, & 0, & ..., & 0, & 0, & 0, & ..., & 0 \\ 0, & \lambda_2, & ..., & 0, & 0, & 0, & ..., & 0 \\ ..., & ..., & ..., & ..., & ..., & ..., & ..., & ... \\ 0, & 0, & ..., & \lambda_M, & 0, & 0, & ..., & 0 \\ 0, & 0, & ..., & 0, & 0, & 0, & ..., & 0 \\ 0, & 0, & ..., & 0, & 0, & 0, & ..., & 0 \\ ..., & ..., & ..., & ..., & ..., & ..., & ..., & ... \\ 0, & 0, & ..., & 0, & 0, & 0, & ..., & 0 \end{bmatrix} \tag{30}$$

- That is, from Equ 29 in conjunction with Equs 27 and 30 we have

$$\mathbb{D} = \underline{\underline{\Lambda}} + \sigma_n^2 \mathbb{I}_N \tag{31}$$

$$\mathbb{D} = \begin{bmatrix} \underbrace{\lambda_1 + \sigma_n^2}_{=d_1} & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & \underbrace{\lambda_2 + \sigma_n^2}_{=d_2} & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \underbrace{\lambda_M + \sigma_n^2}_{=d_M} & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & \underset{=d_{M+1}}{\underset{\uparrow}{\sigma_n^2}} & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \underset{=d_{M+2}}{\underset{\uparrow}{\sigma_n^2}} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & \underset{=d_N}{\underset{\uparrow}{\sigma_n^2}} \end{bmatrix}$$

- This implies that the eigenvalues of the data covariance matrix $\mathbb{R}_{xx}$ (i.e. the diagonal elements of $\mathbb{D}$) are related to the eigenvalues of the emitting signals covariance matrix $\mathbb{R}_{\text{signals}}$ (i.e. diagonal elements of $\underline{\underline{\Lambda}}$) as follows:

$$\begin{aligned} eig_i\left\{\mathbb{R}_{xx}\right\} &= eig_i\left\{\mathbb{R}_{\text{signals}}\right\} + \sigma_{\text{n}}^2 \\ d_i &= \lambda_i + \sigma_{\text{n}}^2 \end{aligned} \quad (32)$$

Now, since the smallest eigenvalue of $\mathbb{R}_{\text{signals}}$ is zero

$$eig_{\min}\left\{\mathbb{R}_{\text{signals}}\right\} = 0 \quad (33)$$

with multiplicity $N-M$, that means

$$eig_{\min}\left\{\mathbb{R}_{xx}\right\} = \sigma_{\text{n}}^{\mathbf{2}} \quad (34)$$

with multiplicity also $N-M$.

- Therefore, theoretically, the number of emitting sources M can be determined by the eigenvalues of the covariance matrix $\mathbb{R}_{xx}$ of the Rx signal-vector $\underline{x}(t)$, and more specifically by the following expression

$$M = N - (\text{multiplicity of } \textbf{minimum} \text{ eigenvalue of } \mathbb{R}_{xx}) \quad (35)$$

- Note: another useful expression is

$$
\begin{aligned}
\mathbb{R}_{\mathbf{xx}} &= \mathbb{E}.\mathbb{D}.\mathbb{E}^H = [\mathbb{E}_{\mathbf{s}}, \mathbb{E}_{\mathbf{n}}].\begin{bmatrix} \mathbb{D}_{\mathbf{s}} & \mathbb{O} \\ \mathbb{O} & \mathbb{D}_{\mathbf{n}} \end{bmatrix}[\mathbb{E}_{\mathbf{s}}, \mathbb{E}_{\mathbf{n}}]^H \\
&= \mathbb{E}_{\mathbf{s}}.\mathbb{D}_{\mathbf{s}}.\mathbb{E}_{\mathbf{s}}^H + \mathbb{E}_n\mathbb{D}_n\mathbb{E}_n^H \qquad (36)
\end{aligned}
$$

where

$$
\mathbb{D} = \begin{bmatrix}
\underbrace{\lambda_1+\sigma_{\mathrm{n}}^2}_{=d_1} & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
0 & \underbrace{\lambda_2+\sigma_{\mathrm{n}}^2}_{=d_2} & \ldots & 0 & 0 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & \underbrace{\lambda_M+\sigma_{\mathrm{n}}^2}_{=d_M} & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & \underset{=d_{M+1}}{\underset{\uparrow}{\sigma_{\mathrm{n}}^2}} & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & \underset{=d_{M+2}}{\underset{\uparrow}{\sigma_{\mathrm{n}}^2}} & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & \underset{=d_N}{\underset{\uparrow}{\sigma_{\mathrm{n}}^2}}
\end{bmatrix}
$$

$\mathbb{D}_{\mathbf{s}}$

$\mathbb{D}_{\mathbf{n}}$

Detection Problem: Finite Observation Interval (L=finite)

Eigen-values

$$\mathbb{D}. = \begin{bmatrix} \underbrace{\lambda_1+\sigma_1^2}_{=d_1} & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & \underbrace{\lambda_2+\sigma_2^2}_{=d_2} & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \underbrace{\lambda_M+\sigma_M^2}_{=d_M} & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & \underset{=d_{M+1}}{\underset{\uparrow}{\sigma_{M+1}^2}} & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \underset{=d_{M+2}}{\underset{\uparrow}{\sigma_{M+2}^2}} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & \underset{=d_N}{\underset{\uparrow}{\sigma_N^2}} \end{bmatrix} \quad (37)$$

N.B.:

- In theory,

$$\sigma_1^2 = \sigma_2^2 = ... = \sigma_M^2 = \sigma_{M+1}^2 = ... = \sigma_N^2 \triangleq \sigma_{\mathrm{n}}^2 \quad (38)$$

- However, in practice

$$\sigma_1^2 \neq \sigma_2^2 \neq ... \neq \sigma_M^2 \neq \sigma_{M+1}^2 \neq .... \neq \sigma_N^2 \quad (39)$$

although

$$\sigma_{\mathrm{n}}^2 \approx \sigma_1^2 \approx \sigma_2^2 \approx ... \approx \sigma_M^2 \approx \sigma_{M+1}^2 \approx .... \approx \sigma_N^2 \quad (40)$$

- if M is **known** then .

$$\begin{aligned} \widehat{\sigma}_{\mathrm{n}}^2 &= \text{the average of the } N-M \text{ smallest eigenvalues} \\ &= \frac{1}{N-M}\left(\sigma_{M+1}^2 + \sigma_{M+2}^2 + ... + \sigma_N^2\right) \end{aligned} \quad (41)$$

- If M is **unknown** then its estimation is not an easy task and a naive approach is likely to fail. Solution: see next section (Decision Theory).

Information Theoretic Criteria

- The information theoretic criteria for model selection, introduced by
  - Akaike (Information Theory Symposium 1973, Automatic Control 1974)
  - Schwartz ("Estimating the dimension of a model," Ann. Stat, 1978), and
  - Rissanen ("Modeling by shortest datad escription," Automatica, 1978)

  address the following general problem.

Problem

Given a set of L observations *(L snapshots), i.e. data,*

$$\mathbb{X} = [\underline{x}(t_1), \underline{x}(t_2)...., \underline{x}(t_L)\}$$

and **a family of models** *, that is, a parameterised family of probability densities pdf($\mathbb{X}|H_i$),$\forall i$, s***elect the model that best fits the data** *(i.e. the model that fits the set of L observations)***.**

- Remember that the conditional probability densities pdf$(\mathbb{X}|H_i)$,$\forall i$, are known as Likelihood Functions (LF)

Solution (AIC)

Select the model which gives the minimum of AIC[a], i.e.

$$\min_k \{AIC(k)\} \tag{42}$$

where

$$AIC(k) \triangleq -2\ln\underbrace{\left(\max_k LF^{(k)}\right)}_{ML\ estimator} + 2k \tag{43}$$

with k denoting the number of free adjusted parameters in the model.

[a] AIC - Akaike Information Criterion

N.B.:

- The first term in Equation-43 is the well-known **log-likelihood of the maximum likelihood** estimator of the parameters of the model.
- The second term is a **bias correction** term, inserted so as to make the AIC an unbiased estimate of the mean Kulback-Liebler distance between the modeled density and the estimated density .

MDL Criterion

- Inspired by Akaike's pioneering work, Schwartz and Rissanen approached the problem from quite different points of view.
    - Schwartz's approach is based on Bayesian arguments (Bayes criterion). He assumed that each competing model can be assigned a prior probability, and proposed to select the model that yields the maximum posterior probability.
    - Rissanen's approach is based on information theoretic arguments. Since each model can be used to encode the observed data, Rissanen proposed to select the model that yields the minimum code length.
- It turns out that in the large-sample limit, both Schwartz's and Rissanen's approaches yield the same criterion, (known as MDL[2] criterion)

[2] MDL: Minimum Description Length

Solution (MDL)

Select the model which gives the minimum of MDL, i.e.

$$\min_k \{MDL(k)\} \tag{44}$$

where

$$MDL(k) \triangleq -\ln \underbrace{\left(\max_k LF^{(k)}\right)}_{ML\ estimator} + \frac{1}{2} k \ln L \tag{45}$$

with k denoting the number of free adjusted parameters in the model.

N.B.:

- Apart from a factor of 2, the **first term** in Equation 45 is identical to the corresponding one in the AIC, while the **second term** has an extra factor of $\frac{1}{2} \ln L$.

Solution (Likehood Function, LF)

It can be proven that

$$LF^{(k)} = -L.\ln\left(\det\left\{\mathbb{R}_{xx}^{(k)}\right\}\right) - Tr\left\{\mathbb{R}_{xx}^{(k)}\right\}^{-1}.\widehat{\mathbb{R}}_{xx} \tag{46}$$

$$= \ln\left(\frac{\prod_{\ell=k+1}^{N} d_\ell^{1/(N-k)}}{\frac{1}{N-k}\sum_{\ell=k+1}^{N} d_\ell}\right)^{(N-k)L} \tag{47}$$

$$\text{where } k \in (0,1,...,N-1)$$

$$\text{and } d_\ell = \text{the } \ell\text{-th eigenvalue of } \widehat{\mathbb{R}}_{xx} \tag{48}$$

- Note that in Equation 47 the term in the brackets is the ratio

$$\frac{\textbf{geometric mean} \text{ of the smallest } N-k \text{ eigenvalues of } \widehat{\mathbb{R}}_{xx}}{\textbf{arithmetic mean} \text{ of the smallest } N-k \text{ eigenvalues of } \widehat{\mathbb{R}}_{xx}} \tag{49}$$

- In Equation 46:
  - $\widehat{\mathbb{R}}_{xx}$ the practical/estimated covariance matrix, i.e,

$$\widehat{\mathbb{R}}_{xx}=\frac{1}{L}\sum_{l=1}^{L}\underline{x}(t_l).\underline{x}^H(t_l)\overset{or}{=}\frac{1}{L}\mathbb{X}.\mathbb{X}^H$$

  - $\mathbb{R}_{xx}^{(k)}$ is the model (theoretical) cov matrix that is used in Equation 21 which defines the N hypotheses, i.e.,

$$\left\{\begin{array}{lll}
H_1: & \mathbb{R}_{xx}^{(0)}=\mathbb{R}_{nn}; & \mathrm{LF}^{(0)}=\mathrm{pdf}_{\underline{x}(t)|H_1}\\
H_2: & \mathbb{R}_{xx}^{(1)}=\underbrace{P_1\underline{S}_1\underline{S}_1^H}_{\mathbb{R}_{signals}}+\mathbb{R}_{nn}; & \mathrm{LF}^{(1)}=\mathrm{pdf}_{\underline{x}(t)|H_2}\\
H_3: & \mathbb{R}_{xx}^{(2)}=\underbrace{\sum_{i=1}^{2}P_i\underline{S}_i\underline{S}_i^H}_{\mathbb{R}_{signals}}+\mathbb{R}_{nn}; & \mathrm{LF}^{(2)}=\mathrm{pdf}_{\underline{x}(t)|H_3}\\
\ldots & \ldots & \ldots\\
H_N: & \mathbb{R}_{xx}^{(N-1)}=\underbrace{\sum_{i=1}^{N-1}P_i\underline{S}_i\underline{S}_i^H}_{\mathbb{R}_{signals}}+\mathbb{R}_{nn}; & \mathrm{LF}^{(N-1)}=\mathrm{pdf}_{\underline{x}(t)|H_N}
\end{array}\right\}$$

Solution (AIC and MDL: equivalent expressions)

It can be proven (based on Equ 47) that

$$AIC(k) = -2\ln\left(\frac{\prod_{\ell=k+1}^{N} d_\ell^{1/(N-k)}}{\frac{1}{N-k}\sum_{\ell=k+1}^{N} d_\ell}\right)^{(N-k)L} + 2k(2N-k) \quad (50)$$

$$MDL(k) = -\ln\left(\frac{\prod_{\ell=k+1}^{N} d_\ell^{1/(N-k)}}{\frac{1}{N-k}\sum_{\ell=k+1}^{N} d_\ell}\right)^{(N-k)L} + \frac{1}{2}k(2N-k)\ln L \quad (51)$$

Reference: M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria", *IEEE Transactions on ASSP*, vol. 33, pp. 387-392, Apr. 1985.

Example

- Consider a ULA of 7 antennas with halfwavelength spacing (i.e. $N = 7$) operating in the presence of two sources with directions-of-arrivals 20° and 25°. The signal-to-noise ratio is 10 dB. Using $L = 100$ snapshots (i.e. samples), the resulted eigenvalues of the sample-covariance matrix are: 21.2359, 2.1717, 1.4279, 1.0979, 1,0544, 0.9432, and 0.7324.
- Observing the gradual decrease of the eigenvalues it is clear that the separation of the "smallest" eigenvalues from the "large" ones is a difficult task.
- However, the AIC and MDL provide the following values

(k)	0	1	2	3	4	5	6
$AIC(k)$	1180.8	100.5	71.4	75.5	86.8	93.2	96.0
$MDL(k)$	590.4	67.2	66.9	80.7	95.5	105.2	110.5
minimum			$M = 2$				

- That is, the minimum of both the AIC and the MDL is obtained, as expected, for the (k) equal to 2, i.e. $M = 2$

AIC Criterion in Vector Format

$$\underline{AIC} = \left[AIC(0),\ \ AIC(1)\ \ ...,\ \ AIC(k),\ \ ...,\ \ AIC(N-1)\right]^T$$

$$= -2L\left(\ln\begin{bmatrix}\prod_{\ell=1}^{N} d_\ell \\ \prod_{\ell=2}^{N} d_\ell \\ \vdots \\ \prod_{\ell=N-2}^{N} d_\ell \\ \prod_{\ell=N-1}^{N} d_\ell \\ d_N\end{bmatrix} + \begin{bmatrix}N \\ N-1 \\ \vdots \\ 3 \\ 2 \\ 1\end{bmatrix} \odot \left(\ln\begin{bmatrix}N \\ N-1 \\ \vdots \\ 3 \\ 2 \\ 1\end{bmatrix} - \ln\begin{bmatrix}\sum_{\ell=1}^{N} d_\ell \\ \sum_{\ell=2}^{N} d_\ell \\ \vdots \\ \sum_{\ell=N-2}^{N} d_\ell \\ \sum_{\ell=N-1}^{N} d_\ell \\ d_N\end{bmatrix}\right)\right)$$

$$+2\begin{bmatrix}0 \\ 1 \\ \vdots \\ N-3 \\ N-2 \\ N-1\end{bmatrix} \odot \begin{bmatrix}2N \\ 2N-1 \\ \vdots \\ N+3 \\ N+2 \\ N+1\end{bmatrix};\quad \text{an } (N\times 1) \text{ real vector} \tag{52}$$

MDL Criterion in Vector Format

$$
\begin{aligned}
\underline{MDL} &= \left[MDL(0),\quad MDL(1)\quad ...,\quad MDL(k),\quad ...,\quad MDL(N-1)\right]^T \\
&= -L\left(\ln\begin{bmatrix} \prod_{\ell=1}^{N} d_\ell \\ \prod_{\ell=2}^{N} d_\ell \\ \vdots \\ \prod_{\ell=N-2}^{N} d_\ell \\ \prod_{\ell=N-1}^{N} d_\ell \\ d_N \end{bmatrix} + \begin{bmatrix} N \\ N-1 \\ \vdots \\ 3 \\ 2 \\ 1 \end{bmatrix} \odot \left(\ln\begin{bmatrix} N \\ N-1 \\ \vdots \\ 3 \\ 2 \\ 1 \end{bmatrix} - \ln\begin{bmatrix} \sum_{\ell=1}^{N} d_\ell \\ \sum_{\ell=2}^{N} d_\ell \\ \vdots \\ \sum_{\ell=N-2}^{N} d_\ell \\ \sum_{\ell=N-1}^{N} d_\ell \\ d_N \end{bmatrix}\right)\right) \\
&\quad +\frac{1}{2}\begin{bmatrix} 0 \\ 1 \\ \vdots \\ N-3 \\ N-2 \\ N-1 \end{bmatrix} \odot \begin{bmatrix} 2N \\ 2N-1 \\ \vdots \\ N+3 \\ N+2 \\ N+1 \end{bmatrix} \ln L; \quad \text{an } (N\times 1) \text{ real vector} \qquad (53)
\end{aligned}
$$

- remember: d_ℓ , (for $\ell = 1$ to N with $d_1 > d_2 > \ldots > d_N$) denotes the ℓ-th eigenvalue of $\mathbb{R}_{xx}$
- Notes:
  - if the **first element** of the vector $\underline{AIC}$ or $\underline{MDL}$ is minimum then $M = 0$
  - if the **second element** of the vector $\underline{AIC}$ or $\underline{MDL}$ is minimum then $M = 1$
  - if the **third element** of the vector $\underline{AIC}$ or $\underline{MDL}$ is minimum then $M = 2$
  - etc.

Detection Problem - Summary

Summary

- If $L=\infty$, i.e. Theoretical $\mathbb{R}_{xx}=\mathcal{E}\left\{\underline{x}(t).\underline{x}(t)^H\right\}$, then

$$\begin{aligned} M &= N-(\text{multiplicity of min. eigenvalue of } \mathbb{R}_{xx}) && (54)\\ \sigma_{\text{n}}^2 &= \text{min.eigenvalue of } \mathbb{R}_{xx} = \text{noise power} && (55) \end{aligned}$$

- If $L=$finite, i.e. Practical $\mathbb{R}_{xx}=\frac{1}{L}\sum_{l=1}^{L}\underline{x}(t_l).\underline{x}(t_l)^H=\frac{1}{L}\mathbb{X}.\mathbb{X}^H$ then

$$\begin{aligned} M &= \text{can be found using AIC or MDL} && (56)\\ \widehat{\sigma}_{\text{n}}^2 &= \text{the average of the } N-M \text{ smallest eigenvalues}\\ &= \frac{1}{N-M}\left(\sigma_{M+1}^2+\sigma_{M+2}^2+....+\sigma_N^2\right) && (57) \end{aligned}$$

- Remember:
  - $N=$number of array elements
  - $M=$ number of signals/sources

The Estimation Problem: Basic Estimation Theory

The Maximum Likelihood (ML) approaches

- Consider an observed $(N \times 1)$ complex signal-vector $\underline{x}(t)$ modelled as follows

$$\underline{x}(t) \triangleq \mathbb{S}(\underline{p}).\underline{m}(t) + \underline{\mathrm{n}}(t) \tag{58}$$

- In this case the L observations at times t_1, t_2,...,t_L (i.e. finite observation interval) are

$$[\underline{x}(t_1), \underline{x}(t_2), \ldots, \underline{x}(t_L)] \text{ defined as the } N \times L \text{ complex matrix } \mathbb{X}$$

since the noise is modelled as a zero mean complex Gaussian random process, with a covariance matrix $\mathbb{R}_{nn} \triangleq \sigma^2 \mathbb{I}_N$ then the observed array signal $\underline{x}(t)$ has a mean vector and covariance matrix which are given as follows:

$$\mathcal{E}\{\underline{x}(t)\} = \mathbb{S}(\underline{p}).\underline{m}(t_l)$$

$$\mathcal{E}\{\underbrace{(\underline{x}(t) - \mathcal{E}\{\underline{x}(t)\}).(\underline{x}(t) - \mathcal{E}\{\underline{x}(t)\})^H}_{\mathbb{R}_{nn}}\} = \sigma_{\mathrm{n}}^2 \mathbb{I}_N$$

- This implies that if there are L observations, which are independent, then the conditional probability density function (likelihood function - LF)

$$\mathrm{LF}_x \triangleq \mathrm{pdf}_x\left(\underline{x}(t_1), \underline{x}(t_2), \ldots, \underline{x}(t_L)\,\middle|\,\underline{p}, \mathbb{M}, \sigma_{\mathrm{n}}^2\right) \tag{59}$$

- ML Solution:

$$\widehat{\mathbb{M}} = (\mathbb{S}(\underline{p})^H \mathbb{S}(\underline{p}))^{-1} \mathbb{S}(\underline{p})^H \mathbb{X} \tag{60}$$

$$\underline{\widehat{m}}(t_i) = \underbrace{(\mathbb{S}(\underline{p})^H \mathbb{S}(\underline{p}))^{-1} \mathbb{S}(\underline{p})^H}_{\triangleq \underline{\mathbb{W}}_{ML}^H} \underline{x}(t_i) \tag{61}$$

$$\begin{aligned} \underline{\widehat{p}}_{ML} &= \arg\max_{\underline{p}} \{\mathrm{LF}_x\} \\ &= \arg\max_{\underline{p}} \left\{ Tr\left(\mathbb{P}_{\mathbb{S}}.\mathbb{R}_{xx}\right)\right\} \end{aligned} \tag{62}$$

where

$$\mathbb{P}_{\mathbb{S}} = \mathbb{S}(\underline{p})\left(\mathbb{S}^H(\underline{p})\mathbb{S}(\underline{p})\right)^{-1}\mathbb{S}(\underline{p})^H \tag{63}$$

The Signal-Subspace Approaches

- In this type of algorithms the parameter M is assumed known $(M < N)$ and involves in some way, or another, two concepts:

  i) the concept of the "**signal-subspace**" associated with the observed signal-vector $\underline{x}(t)$ and its properties

  - ★ This is an *unknown linear subspace* of dimensionality equal M-embedded in an N-dimensional $(M < N)$

  ii) the concept of the "**manifold**" associated with the system/problem's characteristicsin the case of arrays known as "array manifold". It is independent of the noisy observed signal-vector $\underline{x}(t)$ and its properties.

  - ★ This is a *non-linear subspace* (e.g. a curve, surface etc) - embedded in an N-dimensional observation space

- Solution $=$ M points $\begin{cases} \in \text{system's manifold} \\ \in \text{`signal subspace' of } \underline{x}(t) \end{cases}$

- As a result the objective is firstly, from the data, to *estimate the signal subspace* and then to **search the manifold** to find its intersection with the estimated signal-subspace.

The Concept of the "Signal Subspace"

- The first step is to utilize the observed (received) signal vector

$$\underline{x}(t) = \overbrace{\beta_1 m_1(t)}^{\triangleq m_1(t)}\underline{S}_1 + \overbrace{\beta_2 m_2(t)}^{\triangleq m_2(t)}\underline{S}_2 + ... + \overbrace{\beta_M m_M(t)}^{\triangleq m_M(t)}\underline{S}_M + n(t) \quad (64)$$

$$= \mathbb{S}\underline{m}(t) + \underline{\mathrm{n}}(t), \forall t \text{ (infinite observation interval)} \quad (65)$$

or, over L snapshots (finite observation interval)

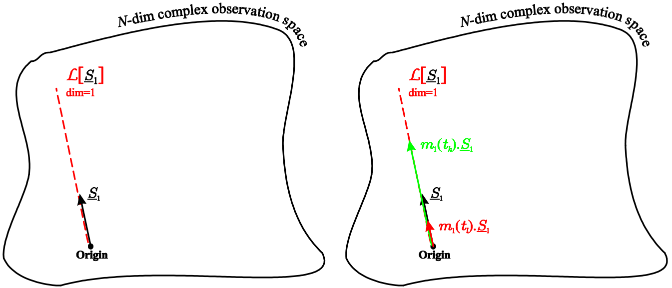
to estimate the "**signal subspace** ".

- The "**signal subspace** " should have **dimensionality** (in most cases) equal to M (known - or estimated) and the signal term $\mathbb{S}.\underline{m}(t)$ belongs always to this subspace.
  i.e.

$$\begin{aligned} &\dim(\textbf{signal subspace}) = M \\ &\mathbb{S}\underline{m}(t) \in \textbf{signal subspace} \end{aligned} \quad (66)$$

The Subspace of a Manifold (Response) Vector

- Let us consider the subspace spanned by only one signal-term of Equation 64, for instance the first term $m_1(t)\underline{S}_1$ at $t = t_1$ and $t = t_2$ as well as the manifold vector $\underline{S}_1$ are shown as below:

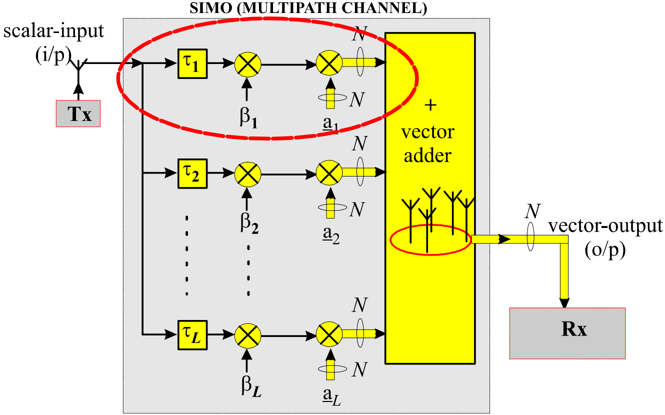


- It is clear from the above figures that

$$\mathcal{L}[\underline{S}_1] = \mathcal{L}[m(t)\underline{S}_1, \forall t] \tag{67}$$

The Subspace of a Manifold (Response) Vector

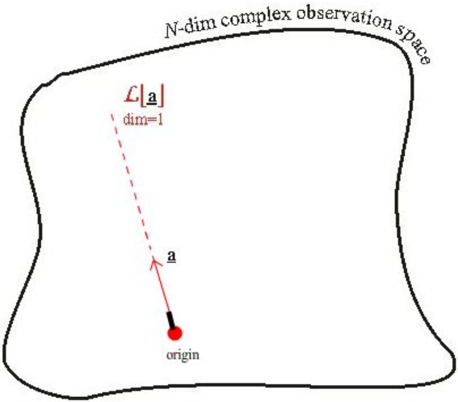
- For instance consider a SIMO multipath channel



- Here the symbols/vectors $\underline{S}$ and $\underline{a}$ are equivalent and both denote an array manifold vector

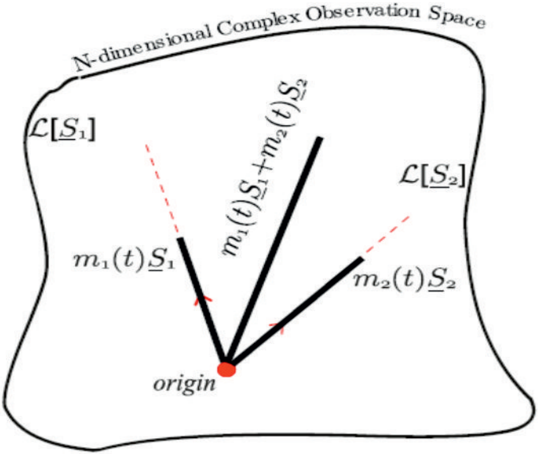
The Subspace of a Single Signal/Source

$$m(t)\underline{a} \Rightarrow \mathcal{L}[\beta m(t)\underline{a}] = \mathcal{L}[\underline{a}] \text{ where } \underline{a} \triangleq \underline{a}(p)$$



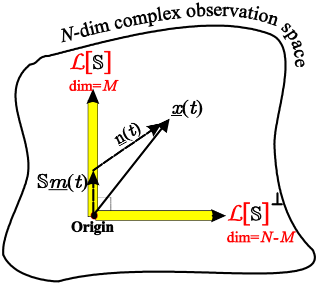
The Subspace of two Signals/Sources

$$m_1(t)\underline{S}_1 + m_2(t)\underline{S}_2 \Rightarrow \mathcal{L}[m_1(t)\underline{S}_1 + m_2(t)\underline{S}_2] = \mathcal{L}[\underline{S}_1, \underline{S}_2]$$

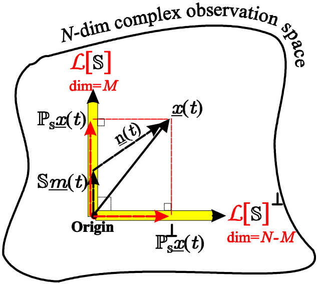


The Subspace of the received signal-vector (t)

- Furthermore, Equation 65 can be represented as follows (at a specific time t):



$\Rightarrow$



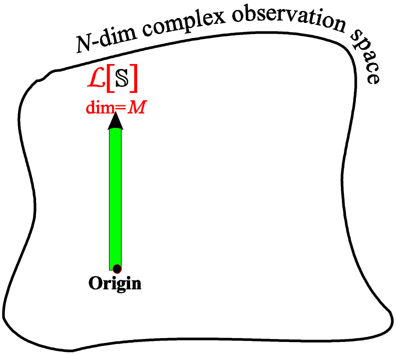
- That is,

$$\begin{aligned}&\text{"subspace of } \underline{x}(t), \forall t\text{"} = \mathcal{L}[\underline{x}(t), \forall t] = \text{whole obs. space} \\ &\text{with } \dim\left(\mathcal{L}[\underline{x}(t), \forall t]\right) = N\end{aligned} \tag{68}$$

- As the dimensionality of this subspace is M and the number of columns of $\mathbb{S}$ is equal to M (remember $\mathbb{S} = [\underline{S}_1, \underline{S}_2, \ldots, \underline{S}_M]$) this implies that

$$\begin{aligned} \text{"signal subspace"} &= \mathcal{L}[\mathbb{S}] \\ \text{with } \dim(\mathcal{L}[\mathbb{S}]) &= M \end{aligned} \tag{69}$$

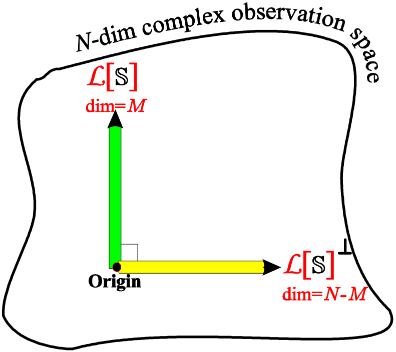
- That is, the signal subspace is spanned by the unknown M manifold vectors associated with the M signals (one signal - one vector)



- The complement subspace to the signal subspace is known as "noise subspace" i.e.

$$\begin{aligned} \text{"noise subspace"} &= \mathcal{L}[\mathbb{S}]^{\perp} \\ \text{with } \dim(\mathcal{L}[\mathbb{S}]^{\perp}) &= N - M \end{aligned}$$

- Signal-Subspace type techniques are based on **partitioning** the observation space into
  - the **Signal Subspace** $\mathcal{L}[\mathbb{S}]$ and
  - the Noise Subspace $\mathcal{L}[\mathbb{S}]^{\perp}$



- However, as the matrix $\mathbb{S}$ remains unknown, the signal subspace and consequently the noise subspace remain unknown.
- Note:

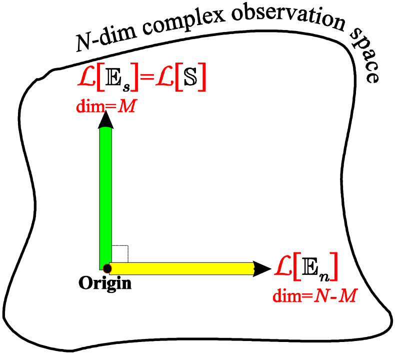
$$\begin{aligned} \text{observation-space} &\triangleq L[\mathbb{X}] \\ &= L[\mathbb{R}_{xx}] \\ &= L[\underline{x}(t), \forall t] \end{aligned} \tag{70}$$

- *Estimation of the two subspaces:* This is achieved by performing an Eigenvector decomposition of the received data covariance matrix $\mathbb{R}_{\mathbf{xx}}$,i.e.

$$\begin{aligned} \mathbb{R}_{\mathbf{xx}} &= \mathbb{E}.\mathbb{D}.\mathbb{E}^H = [\mathbb{E}_{\mathbf{s}}, \mathbb{E}_{\mathbf{n}}].\begin{bmatrix} \mathbb{D}_{\mathbf{s}} & \mathbb{O} \\ \mathbb{O} & \mathbb{D}_{\mathbf{n}} \end{bmatrix} [\mathbb{E}_{\mathbf{s}}, \mathbb{E}_{\mathbf{n}}]^H & (71) \\ \mathbb{R}_{\mathbf{xx}} &= \mathbb{E}_s.\mathbb{D}_s.\mathbb{E}_s^H + \mathbb{E}_n.\mathbb{D}_n.\mathbb{E}_n^H & (72) \end{aligned}$$

- This implies that

$$\boxed{\begin{array}{l|l} \text{signal subspace} = & \mathcal{L}[\mathbb{S}] = \mathcal{L}[\mathbb{E}_s] = \mathcal{L}[\mathbb{E}_n]^\perp \\ \text{noise subspace} = & \mathcal{L}[\mathbb{S}]^\perp = \mathcal{L}[\mathbb{E}_s]^\perp = \mathcal{L}[\mathbb{E}_n] \\ \text{remember:} & \mathbb{S} \neq \mathbb{E}_s \text{ although } \mathcal{L}[\mathbb{S}] = \mathcal{L}[\mathbb{E}_s] \\ \text{observation space} = & \mathcal{L}[\mathbb{R}_{xx}] = \mathcal{L}[\underline{x}(t), \forall t] = \mathcal{L}[\mathbb{X}] \end{array}} \quad (73)$$

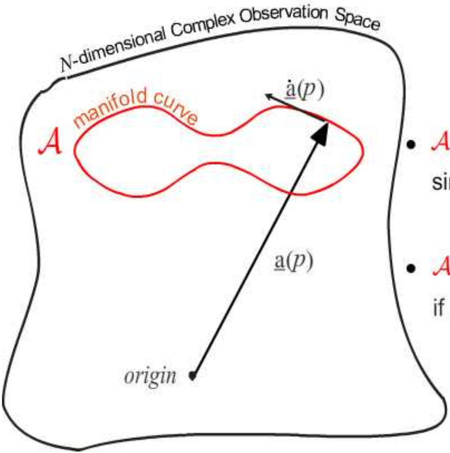

N-dim complex observation space
$\mathcal{L}[\mathbb{E}_s]=\mathcal{L}[\mathbb{S}]$
dim=M
Origin
$\mathcal{L}[\mathbb{E}_n]$
dim=N-M

The Concept of the "Manifold"

One-parameter Manifolds in Wireless Comms

- Consider the manifold vector (or array response vector) of a single parameter p representing for instance θ or ϕ or F_c
  i.e.
  $$\underline{a}(p) \in C^N$$
  This is a single-parameter vector function.

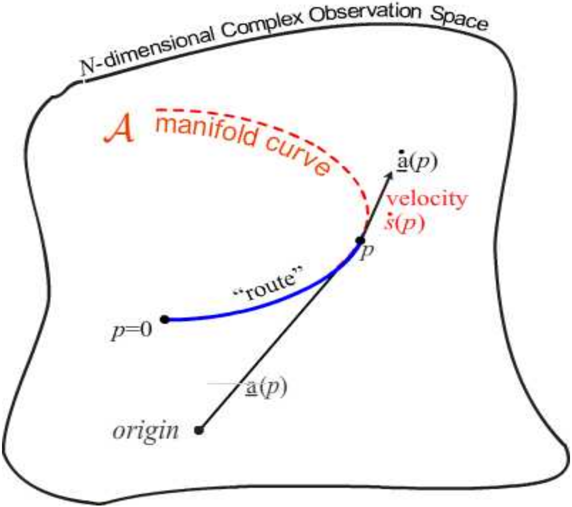
Locus of the Manifold Vectors



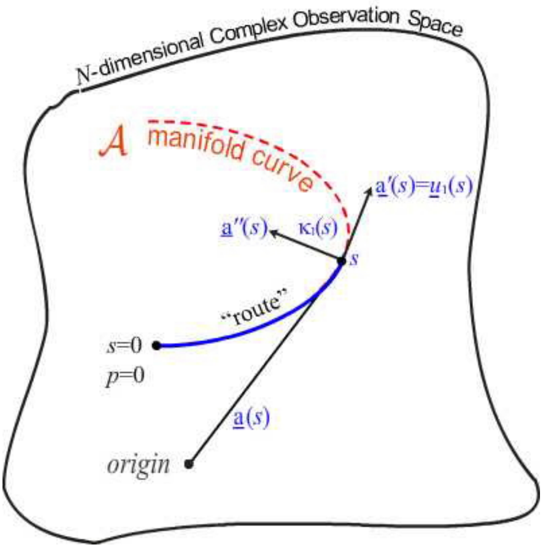
- $\mathcal{A} \triangleq \{\underline{a}(p) \in \mathcal{C}^N, \forall p : p \in \Omega_p\}$ single-parameter curve.
- $\mathcal{A}$ is a regular parameterised curve if $\dot{\underline{a}}(p)$ exists $\forall p$.

- By recording t**he locus of the manifold vectors** as a function of the parameter p (e.g. direction), a "continuum" (i.e. a **geometrical object** such as a curve) is formed **lying in an N-dimensional space**.
- This **geometrical object** (locus of manifold vectors i.e. $\underline{S}(p), \forall p$) is known as **the array manifold**.
- In an array system the manifold (array manifold) can be calculated (and stored) from only the knowledge of the locations and directional characteristics of the sensors.
- Let $\underline{S}(p) \in C^N$ be the manifold vector of a system of N dimensions (e.g. of an array of N sensors) where p is a generic system parameter. This is a single-parameter vector function and as p varies the point $S(p)$ will trace out a curve $\mathcal{A}$ (see figure), embedded in an N-dimensional space C^N.

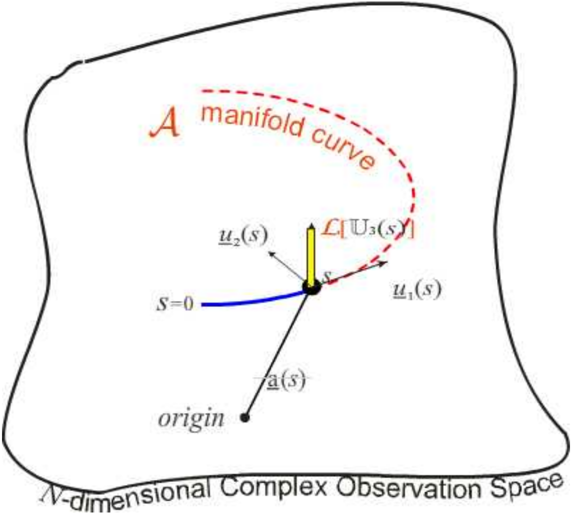
Important Parameters of a Curve



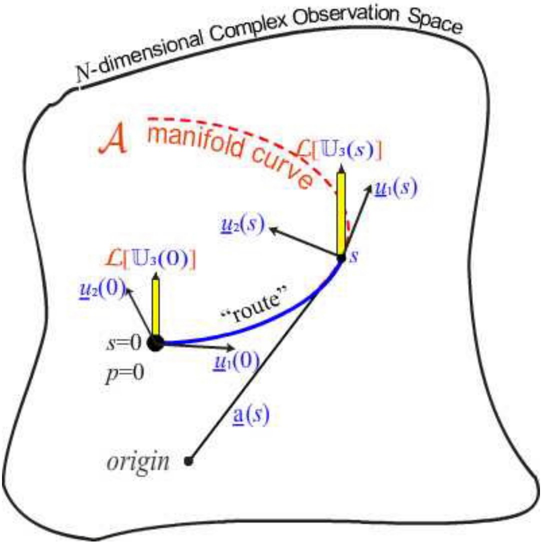
Important Parameters of a Curve



Important Parameters of a Curve

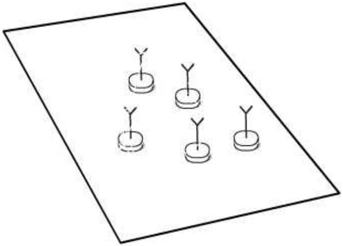


Important Parameters of a Curve

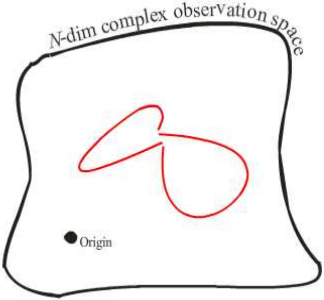


Array of Changing Geometry

- By changing the array geometry the corresponding curve will change

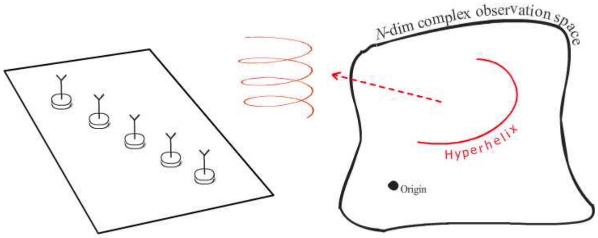


An Array of 5 Sensors with Changing Geometry



Array of Changing Geometry

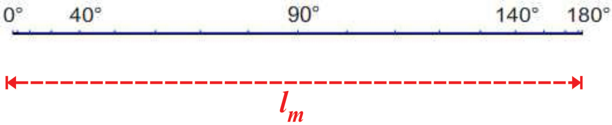
- By changing the array geometry the corresponding curve will change



An Array of 5 Sensors with Changing Geometry

Length of a Curve

- Example:

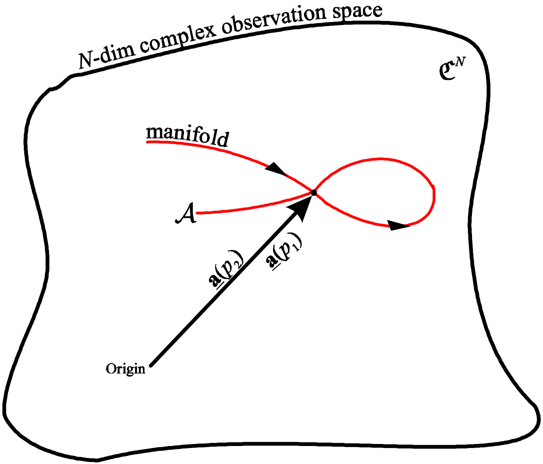


- **important parameters:**
  - **arclength** $s \triangleq s(p) = \int_0^p \left\| \underline{\dot{S}}(p) \right\| dp$
  - **rate-of-change of arclength** $\dot{s} \triangleq \dot{s}(p) = \left\| \underline{\dot{S}}(p) \right\|$
  - **length of manifold** l_m
  - **curvatures** a set of real numbers $\kappa_1, \kappa_2, \kappa_2$, etc (curve's shape)

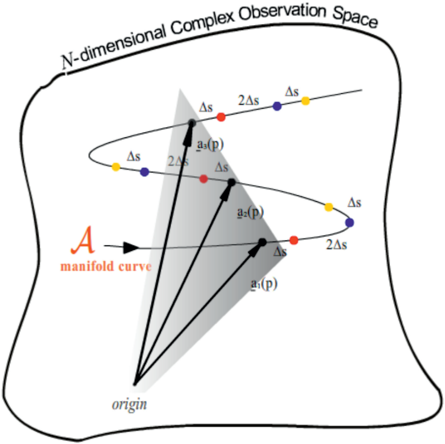
- A curve may have "bad" areas ($\dot{s}$=small) and "good" areas ($\dot{s}$=large)

"Bad" Curves

- There are "bad" and "good" curves (i.e. "bad" and "good" antenna geometries)
- Example of a "bad" curve:

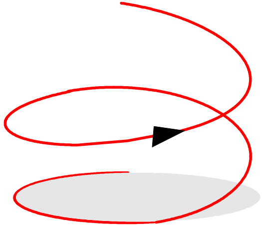


Ambiguities



Single-Parameter Manifolds: Linear Antenna Arrays

- All linear array geometries have manifolds of "hyperhelical" shape embedded in N-dimensional complex space.



Visualisation of a hyperhelix in 3D

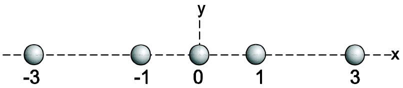
$$\mathbb{C} \triangleq \begin{bmatrix} 0 & -\kappa_1 & 0 & \cdots & 0 & 0 \\ \kappa_1 & 0 & -\kappa_2 & \cdots & 0 & 0 \\ 0 & \kappa_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -\kappa_{d-1} \\ 0 & 0 & 0 & \cdots & \kappa_{d-1} & 0 \end{bmatrix}$$

- Curvatures: forming a matrix known as the **Cartan Matrix** $\mathbb{C}$.
- Hyperhelices: **curvatures=constant** for every s or p
- As we have infinite number of linear arrays we have infinite number of different hyperhelical curves - different set of curvatures

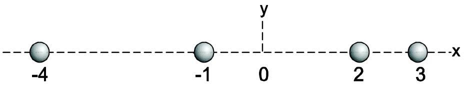
Linear Antenna Array Design

- The Frobenius norm of the Cartan matrix (i.e. $\|\mathbb{C}\|_F$) is related to the array symmetricity as follows:

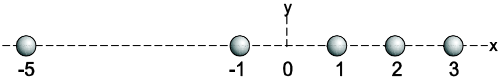
1) $\|\mathbb{C}\|_F = 1$ if array = **symmetric**



2) $\|\mathbb{C}\|_F = \sqrt{2}$ if array = **fully asymmetric**



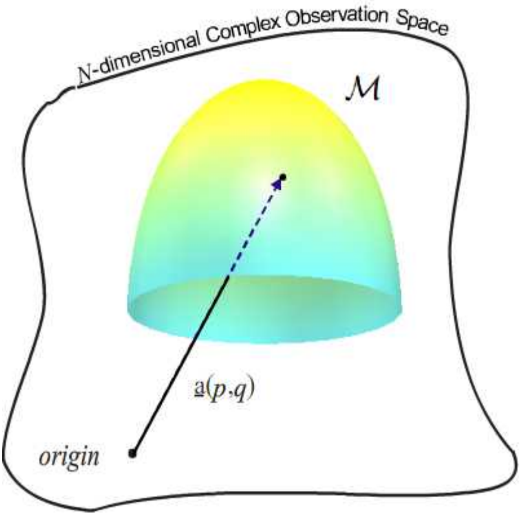
3) $1 < \|\mathbb{C}\|_F < \sqrt{2}$ if array = **partially symmetric**



$\mathrm{eig}(\mathbb{C}) \Rightarrow$ antennas location $\underline{r}_x$

Manifold Surfaces in Wireless Comms

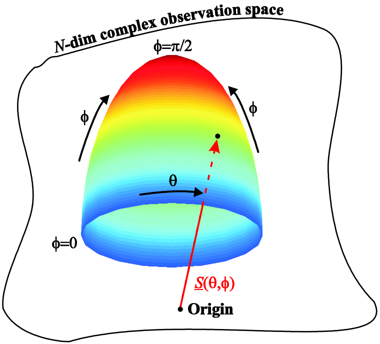
Two-Parameter Manifolds: Visualisation



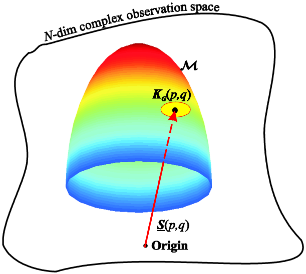
- In a similar fashion *if* there are **two** (unknown) parameters (p,q) per signal *then* $\underline{S}(p,q) \in \mathcal{C}^N$ is a **two-parameter manifold vector** (a vector function)
  and as (p,q) varies the point $\underline{S}(p,q) \in \mathcal{C}^N$ will form a surface $\mathcal{M}$ (see figure), formally defined as follows

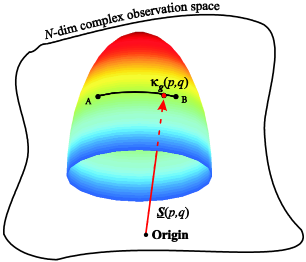
$$\text{Array Manifold: } \mathcal{M} \triangleq \{\underline{S}(p,q) \in C^N, \forall(p,q) : p,q \in \Omega\} \tag{74}$$

where Ω denotes the *parameter space*.



- This surface $\mathcal{M}$ is the locus of all manifold vectors $\underline{S}(p,q)$; $\forall p,q$





- For a point (p, q) on the manifold surface the most important parameters are:
  - *The Gaussian curvature*: $K_G(p, q)$
  - *The manifold metric*: $\mathbb{G}(p, q)$
  - *The Christoffel matrices*: $\underline{\underline{\Gamma}}(p, q)$

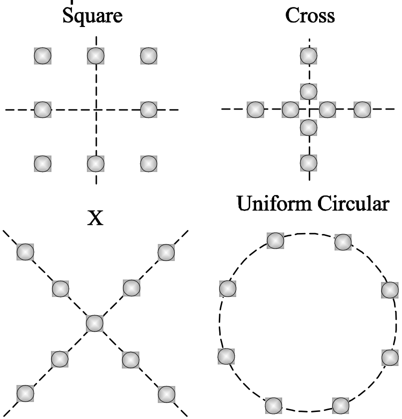
- For a curve on the manifold surface, the parameters of interest are:
  - *The arc length*: s
  - *The geodesic curvature*: κ_g (curve=geodesic$\Rightarrow \kappa_g = 0$)

Some Important Results

- All **planar** antenna geometries (2-Dim arrays) have:

$$\boxed{K_G = 0 \text{ (always)} \Rightarrow \textbf{flat or parabolic of conoid shape} \text{ with the apex at point } \phi{=}90^\circ}$$

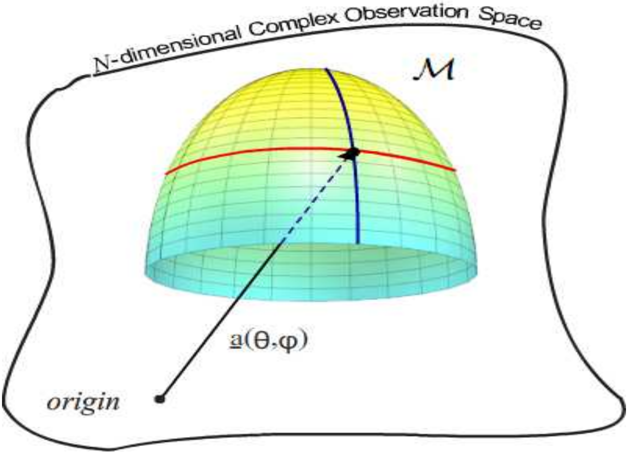
Examples:



- Note: **'flatness' does not imply that there exist straight lines**, as in the case of a surface in $\mathcal{R}^3$.

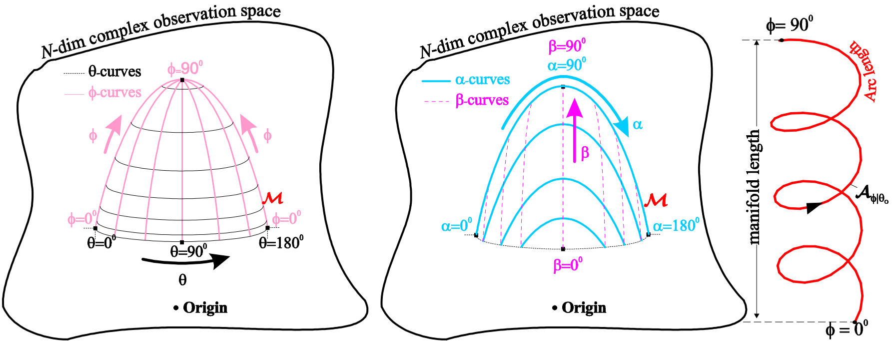
  It means that such surfaces can be **generated** by rotating a **passing geodesic curve** around an apex point.

Manifold Surfaces as Families of Curves

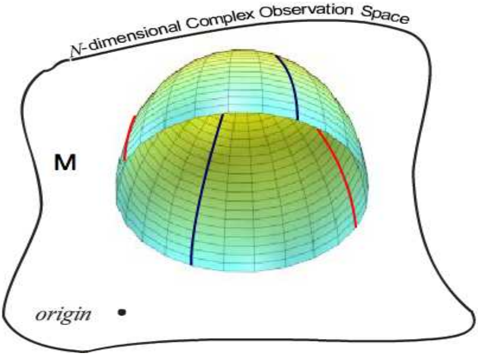


Manifold Surfaces: Parameterisation

- There are **many parameterisations** of a hypersurface.
- We should always try to find a parameterization that makes solving the problem easier.
- e.g.:"Cone"-angles: (α, β) where $\begin{cases} \cos\alpha = \cos\phi\cos(\theta-\Theta) \\ \cos\beta = \cos\phi\sin(\theta-\Theta) \end{cases}$ where Θ =const

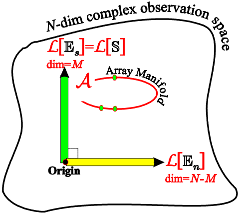


Planar Antenna Array Design Based on Curves

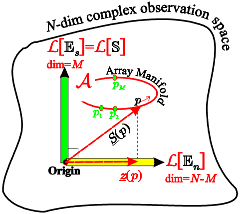


Intersections of Signal Subspace and the Array Manifold

- Both the **manifold** and $\mathcal{L}[\mathbb{E}_s]$ are embedded on the same N-dimensional observation space

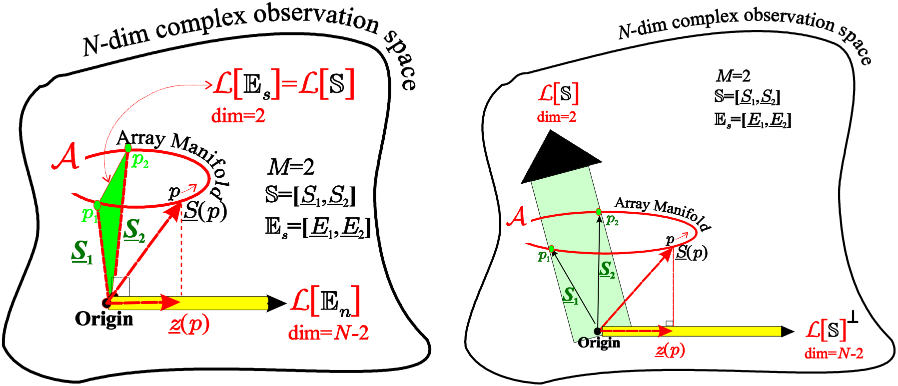


$\Longrightarrow$



- Therefore, the intersection of the manifold with $\mathcal{L}[\mathbb{E}_s]$ will provide the end-points of the columns of the matrix $\mathbb{S}$, i.e. it will provide the parameters $p_1, p_2, \ldots, p_M$.

- Example:



- Note:

$$\underline{S}_1, \underline{S}_2, \underline{E}_1, \underline{E}_2 \quad \perp \quad \mathcal{L}[\mathbb{E}_n]$$
$$\underline{S}_1, \underline{S}_2, \underline{E}_1, \underline{E}_2 \quad \in \quad \mathcal{L}[\mathbb{E}_s]$$

$\mathcal{L}[\mathbb{E}_s]$ is a plane which intersects the array manifold in 2 points $\underline{S}_1, \underline{S}_2$

The MUSIC Algorithm

- The Multiple Signal Classification (MUSIC) algorithm belongs to the family of **Signal-Subspace type technique** .
  i.e.

  $$\boxed{\text{MUSIC} \in \text{Signal-Subspace}}$$

- It estimates the intersection $\mathcal{L}[\mathbb{E}_{\mathbf{s}}]$ and the array manifold by employing the following procedure:

  - Let p denote a parameter value.

  - Form the associated $\underline{S}(p)$ and then project $\underline{S}(p)$ on to the subspace $\mathcal{L}[\mathbb{E}_n]$. This will give us the vector

$$\underline{z}(p) = \mathbb{P}_{\mathbb{E}_n}.\underline{S}(p) \tag{75}$$

- The norm-squared of $\underline{z}(p)$ can be written as

$$\begin{aligned}
\xi(p) &= \underline{z}(p)^H \underline{z}(p) \\
&= \underline{S}(p)^H . \underbrace{\underbrace{\mathbb{P}_{\mathbb{E}_n}^H}_{=\mathbb{P}_{\mathbb{E}_n}} . \mathbb{P}_{\mathbb{E}_n}}_{=\mathbb{P}_{\mathbb{E}_n}} . \underline{S}(p) \\
&= \underline{S}(p)^H . \mathbb{P}_{\mathbb{E}_n} . \underline{S}(p) \\
&= \underline{S}(p)^H . \mathbb{E}_n . (\underbrace{\mathbb{E}_n^H \mathbb{E}_n}_{=\mathbb{I}_{N \times N}})^{-1} . \mathbb{E}_n^H . \underline{S}(p) \\
&= \underline{S}(p)^H . \mathbb{E}_n . \mathbb{E}_n^H . \underline{S}(p)
\end{aligned} \tag{76}$$

- With reference to the figures in slide-78, it is obvious that

$$\xi(p) = 0 \text{ iff } \begin{cases} p = p_1 & \text{or} \\ p = p_2 \end{cases}$$

- Therefore, we search the array manifold, i.e. we evaluate the expression (76),$\forall p$, and we select as our estimates the p's which satisfy

$$\boxed{\xi(p) = 0 \Rightarrow \underline{S}(p)^H . \mathbb{E}_n . \mathbb{E}_n^H . \underline{S}(p) = 0, \forall p} \tag{77}$$

MUSIC Algorithm in Step-format

Step-0: assumptions: M and array geometry are known

Step-1: receive the analogue signal vector $\underline{x}(t) \in C^{N\times 1}$ or its discrete version $\underline{x}(t_l)$, for $l = 1, 2, ..., L$

Step-2: find the covariance matrix

$$\mathbb{R}_{xx} = \begin{cases} \mathcal{E}\left\{\underline{x}(t).\underline{x}(t)^H\right\} \in C^{N\times N} & \text{in theory} \\ \frac{1}{L}\sum_{l=1}^{L}\underline{x}(t_l).\underline{x}(t_l)^H \in C^{N\times N} & \text{in practice} \end{cases}$$

Step-3: find the Eigenvectors and Eigenvalues of $\mathbb{R}_{xx}$

Step-4: form the matrix $\mathbb{E}_s$ with columns the eigevectors which correspond to the M largest eigenvalues of $\mathbb{R}_{xx}$

Step-5: find the arg of the M minima of the function

$$\xi(p) = \underline{S}(p)^H.\mathbb{P}_n.\underline{S}(p), \forall p$$

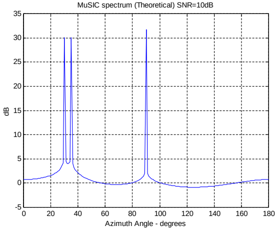
where $\mathbb{P}_n = \mathbb{I}_N - \mathbb{P}_{\mathbb{E}_s}$ and $\mathbb{P}_{\mathbb{E}_s}$ is the projection operator onto $\mathcal{L}\{\mathbb{E}_s\}$,

$$\text{i.e.}\quad [p_1, p_2, .., p_M]^T = \arg\min_p \xi(p)$$

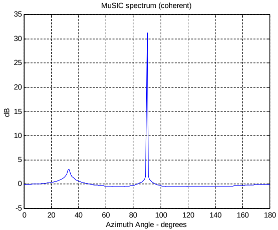
Example (see Experiment AM1)

- Example of MUSIC used in conjunction with a Uniform Linear Array of 5 receiving elements.

- The array operates in the presence of 3 unknown emitting sources with DOA's $(30°, 0°)$, $(35°, 0°)$, $(90°, 0°)$



- MUSIC Limitations:
  - MUSIC breaks down if some incident signals are coherent, i.e. fully correlated, (e.g. multipath situation or 'smart' jamming case)
  - Then $\mathcal{L}[\mathbb{E}_s] \neq \mathcal{L}[\mathbb{S}]$ or, to be more precise, $\mathcal{L}[\mathbb{E}_s] \in \mathcal{L}[\mathbb{S}]$
  - Therefore the 'intersection' argument cannot be used.
  - e.g. same environment as before but the (30°,0°) & (35°, 0°) sources are coherent (fully correlated)



- $\exists$ algorithms which can handle coherent signals in conjunction woth MUSIC: spatial smoothing (see AM1 experiment)

Estimation of Signal Powers, Cross-correlation etc

- Firstly estimate the DOA's and noise power and then use the concept of '**pseudo inverse**' to estimate $\mathbb{R}_{\mathbf{mm}}$
  i.e.

$$\begin{aligned} \text{step} - 1 \quad &: \quad \text{Based on } \mathbb{R}_{xx}, \text{estimate } \underline{p} \text{ and } \sigma_{\mathrm{n}}^2 \\ \text{step} - 2 \quad &: \quad \text{form } \mathbb{S} \\ \text{step} - 3 \quad &: \quad \mathbb{R}_{mm} = \mathbb{S}^{\#} \cdot (\mathbb{R}_{xx} - \sigma_{\mathrm{n}}^2 \mathbb{I}_N) \cdot \mathbb{S}^{\#H} & (78) \\ \text{where } \mathbb{S}^{\#} \quad &= \quad (\mathbb{S}^H \cdot \mathbb{S})^{-1} \cdot \mathbb{S}^H = \text{pseudo-inverse of } \mathbb{S} & (79) \end{aligned}$$

- Note that:
  σ_n^2 =(min eigenvalue of $\mathbb{R}_{xx}$ or given by Equation-41)
  =noise power

Proof.

Proof of Equation-78

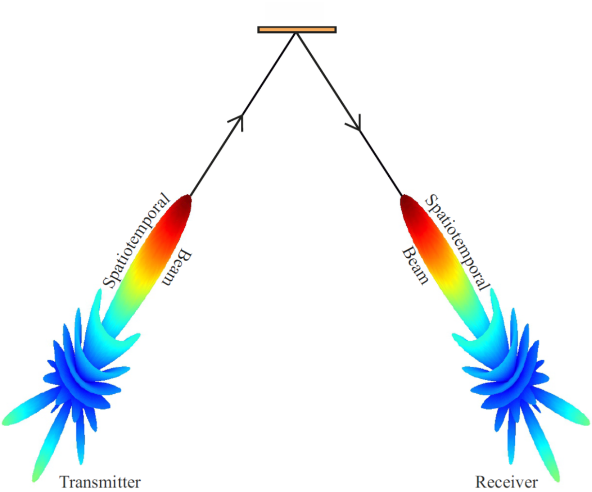
$$\begin{aligned} \mathbb{R}_{xx} &= \mathbb{S}.\mathbb{R}_{mm}.\mathbb{S}^H + \sigma_{\mathrm{n}}^2\mathbb{I}_N \\ \mathbb{R}_{xx} - \sigma_{\mathrm{n}}^2\mathbb{I}_N &= \mathbb{S}.\mathbb{R}_{mm}.\mathbb{S}^H \end{aligned}$$

- By pre & post multiplying both sides of the previous equation with the pseudo inverse of $\mathbb{S}$ we have

$$\begin{aligned} & \mathbb{S}^{\#}.\left(\mathbb{R}_{xx} - \sigma_n^2\mathbb{I}_N\right).\mathbb{S}^{\#H} = \overbrace{\left(\mathbb{S}^H\mathbb{S}\right)^{-1}\mathbb{S}^H}^{=\mathbb{S}^{\#}}.\mathbb{S}\,.\mathbb{R}_{mm}.\mathbb{S}^H.\overbrace{\mathbb{S}\left(\mathbb{S}^H\mathbb{S}\right)^{-1}}^{=\mathbb{S}^{\#H}} \\ \Longrightarrow\quad & \mathbb{S}^{\#}\cdot\left(\mathbb{R}_{xx} - \sigma_n^2\mathbb{I}_N\right)\cdot\mathbb{S}^{\#H} = \mathbb{R}_{mm} \end{aligned}$$

□

The Reception Problem: Array Pattern & Beamforming



[1] This figure is shown in the cover of Prof Manikas' book entitled "Beamforming: Sensor Signal Processing For Defence Applications" .

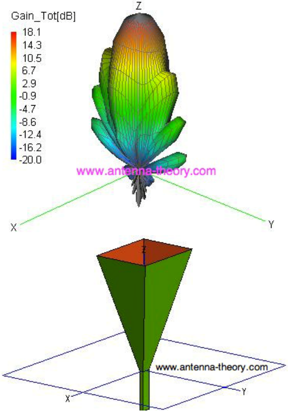
Main Catergories of Beamformers

- **Category-1** : **Single sensor** with directional response.

- Green Bank Telescope, National Radio Astronomy Observatory, West Virginia.
- 100 m clear aperture; Largest fully steerable antenna in the world.

Main Catergories of Beamformers

- **Horn- Antenna** : Another example of **Category-1** (**Single sensor** with directional response).



Main Catergories of Beamformers

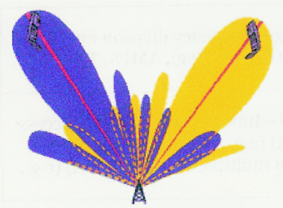
- **Category-2** : **Array of Sensor**
  - Used in SONAR, RADAR, communications, medical imaging, radio astronomy, etc.
  - Line array of directional sensors: Westerbork Synthesis Array Radio Telescope, (WSRT), the Neterlands.

Main Catergories of Beamformers (Array of Sensors)

- **switched beamformer** : there is a finite number of fixed array patterns and the system chooses one of them to maximise signal strength (the one with main lobe closer to the desired user/signal) and switches from one to beam to another as the user/signal moves throughout the sector).
- **adaptive beamformer** (or adaptive array): array patterns are adjusted automatically (main lobe extending towards a user/signal with a null directed towards a cochannel user/signal)

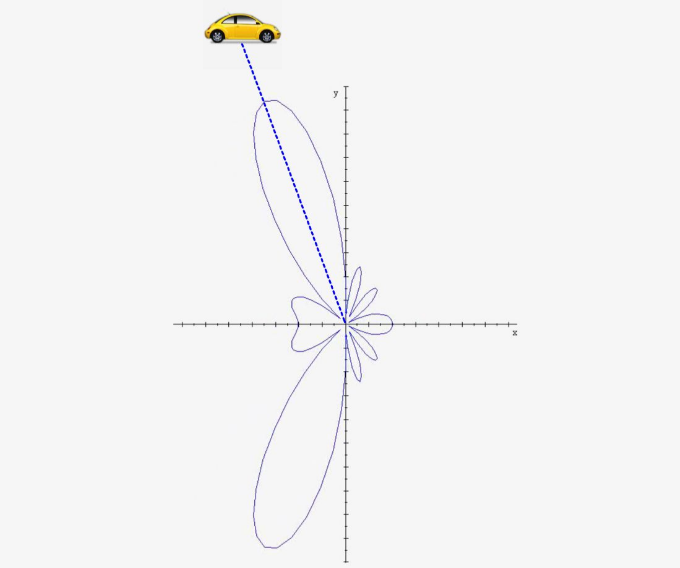
switched beamformer

adaptive beamformer

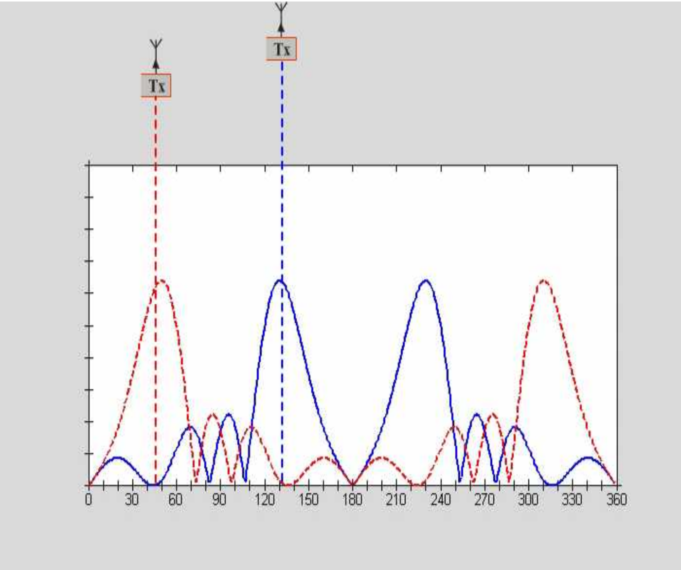


Adaptive for 2 cochannel signals/users

Adaptive Beam (ULA, N=5, d=1)

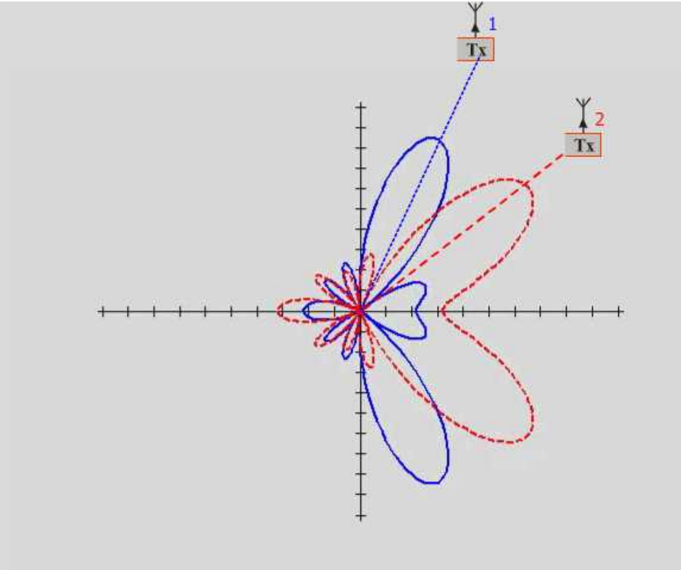


Adaptive Beam (Two co-Channel Signals)



ULA
$N = 5$
$d = 1$

Adaptive Beam (Two co-Channel Signals)



ULA
$N = 5$
$d = 1$

The 'Reception' Problem: Array Pattern & Beamforming

Definitions

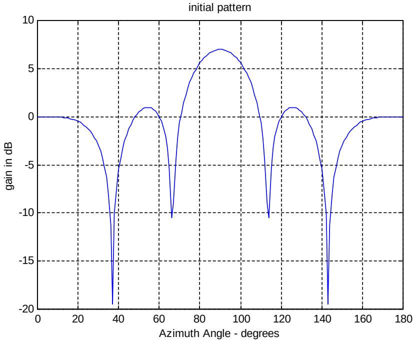
- If the array elements are weighted by complex-weights then the **array pattern** provides the gain of the array as a function of DOAs
  e.g.
  $$\text{if } \theta \longmapsto \underline{S}(\theta) \text{ then } g(\theta) = \underline{w}^H \underline{S}(\theta) \tag{80}$$
  where $g(\theta)$ denotes the gain of the array for a signal arriving from direction θ
  Then,
  $$\boxed{g(\theta), \forall\theta : \textbf{is known as the array pattern}} \tag{81}$$

The array pattern is a function of the array manifold $\underline{S}(\theta)$ (i.e. array geometry and channel parameter θ) and the Rx weight vector $\underline{w}$

- N.B.: default pattern:

$$g(\theta) = \underline{1}_N^T \underline{S}(\theta) \tag{82}$$

i.e. $\underline{w} = \underline{1}_N$ (i.e. no weights)
e.g. Array Pattern of a Uniform Linear Array of 5 elements
$(\underline{w} = \underline{1}_5$, i.e. no weights)



- **Beamwidth**
  The array pattern **has** a number of **lobes.**
  - the **largest** lobe is called the **'main lobe'** while
  - the **remaining** lobes are known as **'sidelobes'** .

$$\text{beamwidth}^{\circ} = 2\sin^{-1}\left(\frac{\lambda}{Nd}\right) \times \frac{180}{\pi} \tag{83}$$

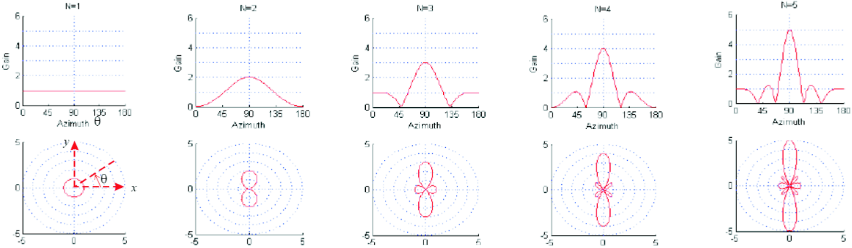
- Note that d =inter-sensor-spacing, and

$$\text{if } d=\frac{\lambda}{2} \Rightarrow \text{beamwidth}^{\circ}=2\sin^{-1}\left(\frac{2}{N}\right)\times\frac{180}{\pi} \tag{84}$$

- To **steer the main lobe** towards a specific (known) direction θ, a **'spatial correction weight'** $\underline{w}_{\text{main-lobe}}$ can be used which should be equal to

$$\begin{aligned} \underline{w}_{\text{main-lobe}} &= \exp\left(-j\underline{\mathbf{r}}^{T}\underline{k}_{\text{main-lobe}}\right) && (85)\\ \underline{w}_{\text{main-lobe}} &= S\left(\theta_{\text{main-lobe}}\right) && (86) \end{aligned}$$

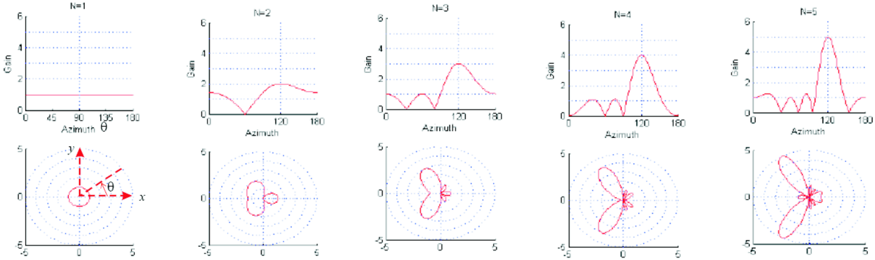
- ARRAY PATTERN: for arrays of $N = 1, 2, 3, 4$ and 5 sensors (Mainlobe at 90°, $d = 1$ half-wavelength)



$$\underline{r}_x = [0] \quad = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}^T \quad = \begin{bmatrix} -1 \\ 0 \\ +1 \end{bmatrix}^T \quad = \begin{bmatrix} -1.5 \\ -0.5 \\ 0.5 \\ 1.5 \end{bmatrix}^T \quad = \begin{bmatrix} -1.5 \\ -0.5 \\ 0 \\ 0.5 \\ 1.5 \end{bmatrix}^T$$

$$\underline{w} \rightarrow = \quad = [1,1]^T \quad = [1,1,1]^T \quad = [1,1,1,1]^T \quad = [1,1,1,1,1]^T$$

- ARRAY PATTERN for ULA arrays of $N = 2, 3, 4, 5$ sensors (mainlobe at $120°$, $d = 1$ half-wavelength)

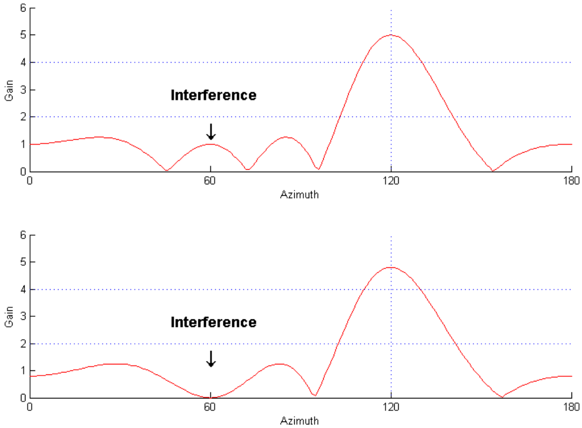


$$\underline{r}_x = [0] \quad = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}^T \quad = \begin{bmatrix} -1 \\ 0 \\ +1 \end{bmatrix}^T \quad = \begin{bmatrix} -1.5 \\ -0.5 \\ 0.5 \\ 1.5 \end{bmatrix}^T \quad = \begin{bmatrix} -1.5 \\ -0.5 \\ 0 \\ 0.5 \\ 1.5 \end{bmatrix}^T$$

$$\underline{w} = \underline{S}(120°, 0) = \exp(-j\underline{r}^T \underline{k}(120°, 0°)) \quad (87)$$

$$(\text{simplified to}) = \exp(-j\pi \underline{r}_x \cos(120°)) \quad (88)$$

- A **beamformer** is an array system which **receives** a 'desired' signal and **suppresses** (according to a criterion) co-channel interference and noise effects, by **synthesizing an array pattern** with high-gain towards the DOA of the desired signal and deep nulls towards the DOAs of the interfering signals (adaptive arrays).



Some Popular Beamformers

- WIENER-HOPF Beamformer:

$$\begin{aligned} \underline{w} &= c.\mathbb{R}_{xx}^{-1}\underline{S}_{\substack{\text{desired}\\ \text{signal}}} \qquad (89) \\ \text{where } c &= \text{a constant scalar} \end{aligned}$$

  - Maximizes the SNIR at the array output.
  - It is optimum wrt SNIR criterion
  - It is a conventional beamformer (i.e. resolution is a function of the SNR_{in})
  - No need to know the DOAs of the interfering signals
  - (please try to prove Equation-89)

- Modified WIENER-HOPF Beamformer:

$$\begin{aligned} \underline{w} &= c \cdot \mathbb{R}_{n+J}^{-1} \underline{S}_{\substack{\text{desired}\\ \text{signal}}} \\ \text{where } c &= \text{a constant scalar} \end{aligned} \tag{90}$$

  - comments similar to Wiener-Hopf
  - robust to ‘pointing’ errors (i.e. robust to errors associated with the direction of the desired signal)

- Minimum Variance Beamformer
  - It is also known as Capon's beamformer
  - It is the beamformer that solves the following optimisation problem

$$\min_{\underline{w}} \left(\underline{w}^H \mathbb{R}_{xx} \underline{w} \right) \tag{91}$$

$$\text{subject to } \underline{w}^H \underline{S}(\theta) = 1 \tag{92}$$

  - Equation 91 aims to minimise the effect of the desired signal plus noise. However, the constraint $\underline{w}^H \underline{S}(\theta) = 1$ (Equ 92) prevents the gain reduction in the direction of the desired signal.
  - Solution of Equations 91 and 92:

$$\underline{w} = c.\mathbb{R}_{xx}^{-1} \underline{S}_{\substack{\text{desired}\\ \text{signal}}} \tag{93}$$

$$\text{where } c = \text{a constant scalar chosen such as } \underline{w}^H \underline{S}(\theta) = 1$$

- A Superresolution Beamformer based on DOA estimation:

$$\begin{aligned} \underline{w} &= \mathbb{P}^{\perp}_{\mathbb{S}_J} \underline{S}_{\substack{\text{desired}\\\text{signal}}} \\ \text{where } \mathbb{S} &= \left[\underline{S}_{\substack{\text{desired}\\\text{signal}}}, \mathbb{S}_J\right] \end{aligned} \tag{94}$$

  - Provides complete (asymptotically) interference cancellation.
  - Maximizes the SIR at the array output.
  - It is optimum wrt SIR criterion
  - It is a superresolution beamformer (i.e. resolution is not a function of the SNR_{in})
  - Needs an estimation algorithm to provide the DOAs of all incident signals.

- A Superresolution Beamformer not based on DOA estimation of interfering sources

$$\underline{w} = \mathbb{P}^{\perp}_{\mathbb{E}_{n_j}} \cdot \underline{S}_{\substack{\text{desired}\\ \text{signal}}} \tag{95}$$

where $\mathbb{E}_{n_j}$ =noise subspace of $\mathbb{R}_{n+J}$
Note: $\mathbb{R}_{n+J}$ =covariance matrix where the effects of the

desired signal have been removed

- Maximum Likelihood (ML) Beamformer

$$\underline{w} = \mathrm{col}_{\text{des}}\left(\mathbb{S}^{\#}\right) \tag{96}$$

$$\text{where } \mathbb{S}^{\#} = \mathbb{S}.(\mathbb{S}^{H}\mathbb{S})^{-1} \tag{97}$$

Examples of Array Patterns (Beamformers)

- Consider a uniform linear array of 5 elements operating in the presence of three signals with directions $(90^\circ, 0^\circ)$, $(30^\circ, 0^\circ)$ and $(35^\circ, 0^\circ)$. One of the signals is the 'desired' signal and the other two are unknown interferences.

- Initially, the 'desired' DOA is

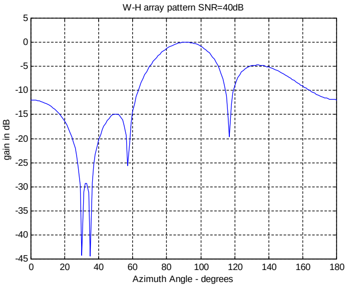
$$(90^\circ, 0^\circ) = \text{known}$$

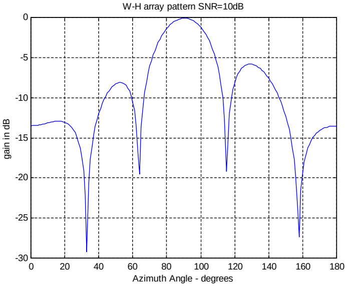
and the DOAs of interfering sources are:

$$\begin{aligned}(30^\circ, 0^\circ) &= \text{unknown}\\ (35^\circ, 0^\circ) &= \text{unknown}\end{aligned}$$

WIENER-HOPF Beamformer (Equ. 89):

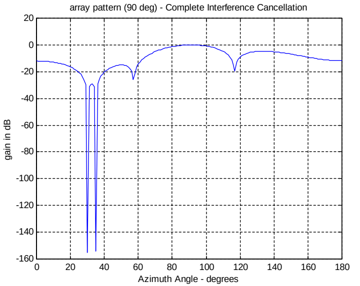
SNR=40dB (high) or 10dB (low)





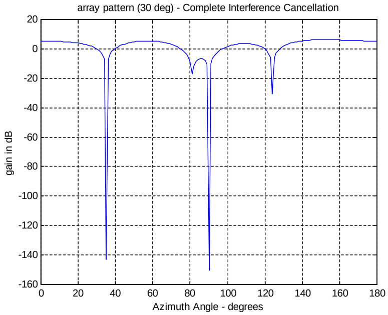
Superresolution Beamformer (Equ. 95)

SNR=10dB and 40dB

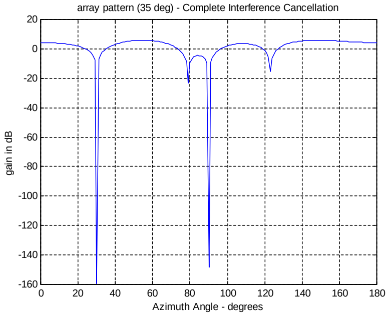


- Superresolution Beamformer (Equ.-94) (all DOAs known).

a) desired source=30°



b) desired source=35°



Beamformers in Mobile Communications

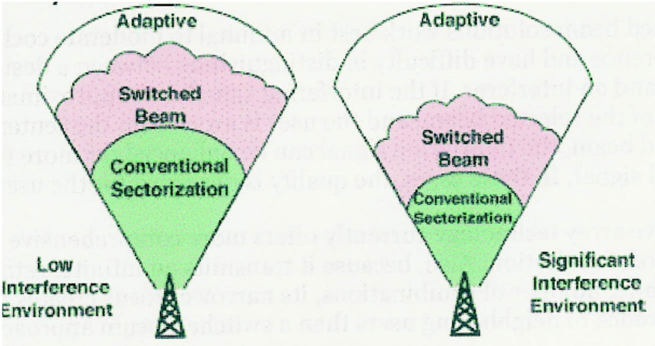
- Some Applications of Beamformers in Communications:

1	analogue access methods	FDMA (e.g. AMPS, TACS, NMT)
2	digital access methods	TDMA (e.g GSM, IS136), CDMA
3	duplex methods	FDD, TDD

- Main properties of beamforming in Communications

	Properties	Advantages
1	signal gain	better range/coverage(see figure below)
2	interference rejection	increase capacity
3	spatial diversity	multipath rejection
4	power efficiency	reduced expense

- Coverage patterns for switched beam and adaptive array antenna



Array Performance Criteria and Bounds

Introduction

- Two popular performance evaluation criteria are:
  - SNIR_{out}
  - Outage Probability
- Three Array Bounds
  - Detection Bound
  - Resolution Bound
  - Cramer-Rao Bound (estimation accuracy)

SNIRout Criterion

- The signal at the output of the beamformer can be expressed as

$$\begin{aligned} y(t) &= \underline{w}^H \underline{x}(t) = \underline{w}^H (\mathbb{S}\underline{m}(t) + \underline{n}(t)) \\ &= \underline{w}^H (\underline{S}_1 m_1(t) + \mathbb{S}_J \underline{m}_J(t) + \underline{n}(t)) \end{aligned}$$

where

$$\begin{aligned} \mathbb{S} &= [\underline{S}_1, \underbrace{\underline{S}_2, \ldots, \underline{S}_M}_{\triangleq \mathbb{S}_J}] \\ \underline{m}(t) &= [m_1(t), \underbrace{m_2(t), \ldots, m_M(t)}_{\triangleq \underline{m}_J^T}]^T \end{aligned}$$

- Power of $y(t)$:

$$
\begin{aligned}
P_y &= \mathcal{E}\left\{y(t)^2\right\} = \\
&= \mathcal{E}\left\{y(t)y(t)^*\right\} = \mathcal{E}\left\{\underline{w}^H \underline{x}(t)\underline{x}(t)^H \underline{w}\right\} \\
&= \underline{w}^H \underbrace{\mathcal{E}\left\{\underline{x}(t)\underline{x}(t)^H\right\}}_{=\mathbb{R}_{xx}} \underline{w} \\
&= \underline{w}^H \left(\underbrace{P_1 \underline{S}_1 \underline{S}_1^H}_{\triangleq \mathbb{R}_{dd}} + \underbrace{\mathbb{S}_J \mathbb{R}_{m_J m_J} \mathbb{S}_J^H}_{\triangleq \mathbb{R}_{JJ}} + \underbrace{\sigma^2 \mathbb{I}_N}_{\triangleq \mathbb{R}_{nn}} \right) \underline{w} \\
&= \underline{w}^H (\mathbb{R}_{dd} + \mathbb{R}_{JJ} + \mathbb{R}_{nn}) \underline{w} \\
&\quad \text{(assuming desired, interfs \& noise are uncorrelated)}
\end{aligned}
$$

- i.e.

$$P_y = \underbrace{\underline{w}^H \mathbb{R}_{dd} \underline{w}}_{=P_{d,out}} + \underbrace{\underline{w}^H \mathbb{R}_{JJ} \underline{w}}_{=P_{J,out}} + \underbrace{\underline{w}^H \mathbb{R}_{nn} \underline{w}}_{=P_{n,out}}$$

where

$$\begin{aligned} P_{d,out} &= \text{o/p desired term} \\ &= P_1 \underline{w}^H \underline{S}_1 \underline{S}_1^H \underline{w} = P_1 (\underline{w}^H \underline{S}_1)^2 \end{aligned}$$

$$\begin{aligned} P_{J,out} &= \text{o/p interf. term} \\ &= \sum_{i=2}^{M} \sum_{j=2}^{M} \rho_{ij} \underline{w}^H \underline{S}_i \underline{S}_j^H \underline{w} \text{ with } \rho_{ij} \triangleq \text{corr.coeff} \end{aligned}$$

$$\begin{aligned} P_{n,out} &= \text{o/p noise term} \\ &= \sigma_{\mathrm{n}}^2 \underline{w}^H \underline{w} \end{aligned}$$

- Therefore,

$$\mathrm{SNIR}_{\mathrm{out}} = \frac{P_{d,out}}{P_{J,out} + P_{n,out}} = \frac{P_1 \left(\underline{w}^H \underline{S}_1 \right)^2}{\underbrace{\sum_{i=2}^{M} \sum_{j=2}^{M} \rho_{ij} \underline{w}^H \underline{S}_i \underline{S}_j^H \underline{w}}_{\underline{w}^H \mathbb{S}_J \mathbb{R}_{m_J m_J} \mathbb{S}_J^H \underline{w}} + \sigma_n^2 \underline{w}^H \underline{w}} \quad (98)$$

- Note:
  - An alternative equivalent expression to Equ 98 is given below

$$\mathrm{SNIR}_{\mathrm{out}} = \frac{\underline{w}^H \mathbb{R}_{dd} \underline{w}}{\underline{w}^H \left(\mathbb{R}_{xx} - \mathbb{R}_{dd} \right) \underline{w}} \quad (99)$$

  - Both Equations 98 and 99 are general expressions for any beamformer. However, for different beamformers (i.e. different weights) these equations give different values.

Outage Probability Criterion

- outage probability (OP) is defined as follows:

$$\text{OP} = \Pr(\text{SNIR}_{\text{out}} < \text{SNIR}_{pr}) \tag{100}$$

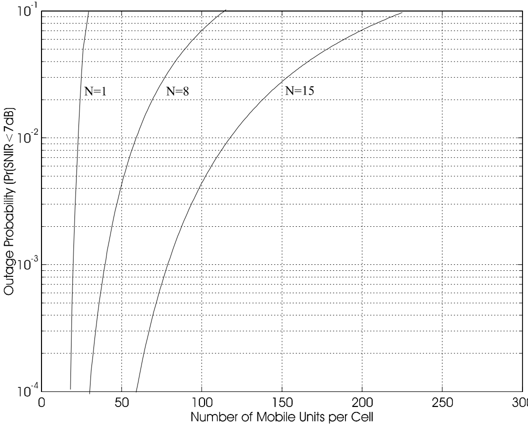
or

$$\text{OP} = \Pr(\text{SIR}_{\text{out}} < \text{SIR}_{pr}) \tag{101}$$

- It is a performance evaluation criterion.

- An example of an array-CDMA system's Outage Probability with $N = 1, 8, 15$ receiving elements is shown below clearly showing that by employing an antenna array and using, for instance, the beamformer of Equation 94 (complete interference cancellation beamformer) at the base station, the system capacity is increased.

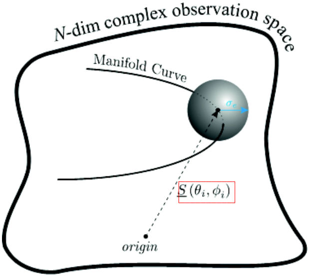
- Outage Probability Examples/Graphs



- For example, for 0.001 outage probability, the system capacity per cell increases from 20 mobiles, for a single antenna case (i.e. $N = 1$), to about 40 and 80 mobiles for N equal to 8 and 15, respectively.

Uncertainty Hyperspheres and CRB

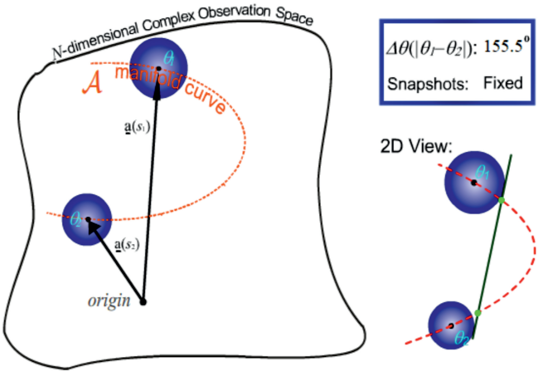
Model the uncertainty remaining in the system after L snapshots as an *Uncertainty Hypersphere* of effective radius σ_e



$$
\begin{aligned}
\sigma_e &= \sqrt{\mathrm{CRB}[s]} \qquad (102) \\
&= \sqrt{\frac{1}{2\,(\mathrm{SNR} \times L)\, C}} \\
\mathrm{CRB} &= \text{Cramer-Rao Bound}
\end{aligned}
$$

- [see chapter-8 of my book]: The uncertainty sphere represents the smallest achievable uncertainty (optimal accuracy) due to the presence of noise after L snapshots, when all the effects of the presence of other sources have been eliminated by an "ideal" parameter estimation algorithm
- The Parameter C ($0 < C \leq 1$): models any additional uncertainty introduced by a practical parameter estimation algorithm. **N.B.: Ideal algorithm:** $C = 1$

Detection and Resolution Bounds



Detection and Resolution Bounds

Angular Separation: Resolution & Detection Laws

- Note: $\sigma_e \propto \sqrt[-2]{SNR \times L}$; where L =number of snapshots
- **Square-root Law** :

$$\text{Detection: } \Delta p = \mathsf{f}\left\{\dot{s}, \sigma_e\right\} \tag{103}$$

- **Fourth-root Law** :

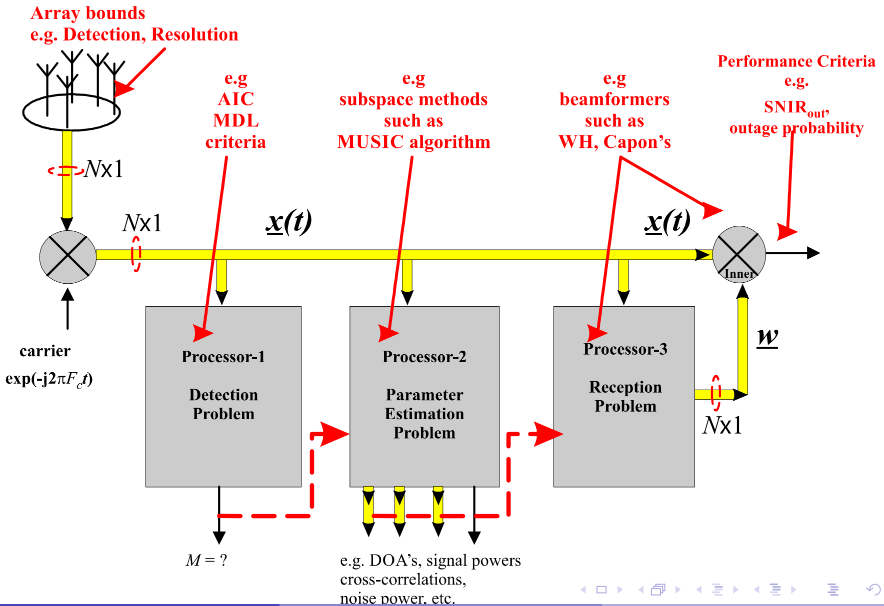
$$\text{Resolution: } \Delta p_{th} = \mathsf{f}\left\{\dot{s}, \sqrt[4]{\kappa_1}, \sqrt{\sigma_e}\right\} \tag{104}$$

- Remember - **Frequency Selective Channels** :
  - number of resolvable paths$=\left\lfloor \frac{\textit{Delay-Spread}}{\text{channel-symbol period}} \right\rfloor + 1$

    $\Downarrow$
  - two paths with a relative delay $<$ channel-symbol-period

    cannot be resolved
- for more info see Chapter 8 of my book

Summary



Array bounds
e.g. Detection, Resolution
e.g
AIC
MDL
criteria
e.g
subspace methods
such as
MUSIC algorithm
e.g
beamformers
such as
WH, Capon's
Performance Criteria
e.g.
SNIR$_{out}$,
outage probability
Nx1
Nx1
x(t)
x(t)
Inner
w
carrier
exp(-j2πF$_c$t)
Processor-1
Detection
Problem
Processor-2
Parameter
Estimation
Problem
Processor-3
Reception
Problem
Nx1
M = ?
e.g. DOA's, signal powers
cross-correlations,
noise power, etc.

Appendix-A: Basic Decision Theory

Decision Criteria

Consider an observed signal $r(t)$ and M hypotheses H_1, H_2,..., H_M.

- Define the sets of parameters $\mathcal{P}_1$ and $\mathcal{P}_2$ where:

  $\mathcal{P}_1$ denotes the set of parameters $\Pr(H_1)$,$\Pr(H_2)$, ..., $\Pr(H_M)$
  $\mathcal{P}_2$ represents the set of costs/weights C_{ij}, $\forall i, j$
  associated with the probabilities $\Pr(D_i|H_j)$,
  (where D_i indicates "decision" by choosing hypothesis H_i)

- Estimate/identify the likelihood functions

$$\underbrace{\mathrm{pdf}_{r|H_1}(r)}_{\substack{\updownarrow \\ LF^{(1)}}}, \underbrace{\mathrm{pdf}_{r|H_2}(r)}_{\substack{\updownarrow \\ LF^{(2)}}}, \ldots, \underbrace{\mathrm{pdf}_{r|H_M}(r)}_{\substack{\updownarrow \\ LF^{(M)}}}$$

Main Decision Criteria

Decision: choose the hypothesis H_i (i.e. D_i) with the maximum $G_i(r)$ where $G_i(r)$ depends on the chosen criterion as follows:

- **BAYES** Criterion
- **Minimum Probability of Error** $(\min(\mathrm{p_e}))$ Criterion
- **MAP** Criterion
- **MINIMAX** Criterion
- Newman-Pearson (**N-P**) Criterion
- Maximum Likelihood (**ML**) Criterion

	$\mathcal{P}_1$	$\mathcal{P}_2$	choose Hypothesis with $\max(G_j(r))$
Bayes	**known**	**known**	$G_i(r) \triangleq \textbf{weight}_i \times \Pr(H_i) \times \mathrm{pdf}_{r\mid H_i}$
$\min(p_e)$ or MAP	**known**	unknown	$G_i(r) \triangleq \Pr(H_i) \times \mathrm{pdf}_{r\mid H_i}$
Minimax	unknown	**known**	$G_i(r) \triangleq \textbf{weight}_i \times \mathrm{pdf}_{r\mid H_i}$
N-P	unknown	unknown	by solving a constraint optim. problem
ML	don't care	don't care	$G_i(r) \triangleq \mathrm{pdf}_{r\mid H_i}$

- N.B.:
  - Note-1:
    if an approximate/initial solution is required then any information about the sets of parameters $\mathcal{P}_1$ and/or $\mathcal{P}_2$ can be ignored. In this case the Maximum Likelihood (ML) Criterion should be used.
  - Note-2:

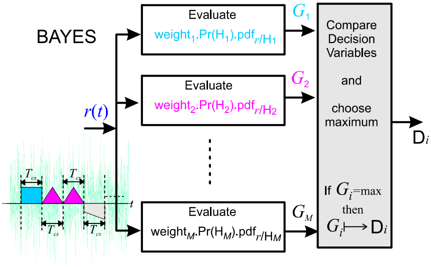
$$\text{weight}_i \triangleq \sum_{\substack{j=1 \\ j \neq i}}^{M} \left(C_{ji} - C_{ii} \right) \tag{105}$$

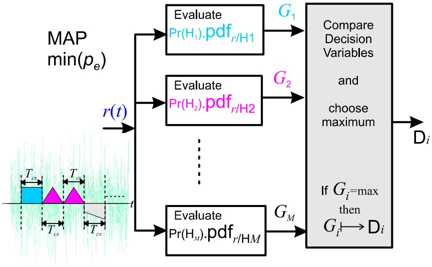
    or (since the term for $i = j$ is equal to zero) simply,

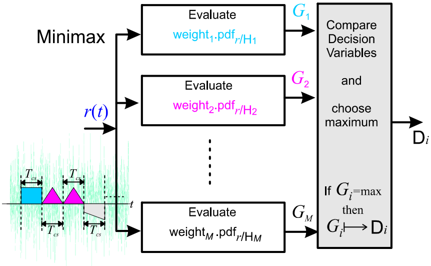
$$\text{weight}_i = \sum_{j=1}^{M} \left(C_{ji} - C_{ii} \right) \tag{106}$$

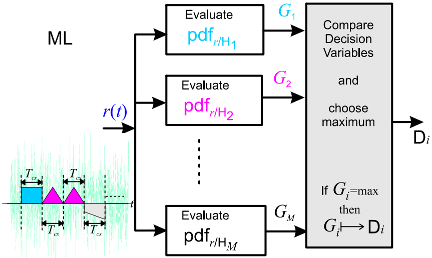
  - Note-3:
    Sometimes, for convenience, G_i will be used (i.e. $G_i \triangleq G_i(r)$) - i.e. the argument will be ignored.
  - Note-4:
    for binary: $\text{weight}_0 \triangleq C_{10} - C_{00}$; $\text{weight}_1 \triangleq C_{01} - C_{11}$

Decision Criteria: Mathematical Architectures









Appendix-B: Optimum M-ary Rx's and Decision Theory

- **Objective** : to design a receiver which operates on $r(t)$ and chooses one of the following M hypotheses:

$$\begin{cases} H_1: & r(t) = s_1(t) + n(t) \\ H_2: & r(t) = s_2(t) + n(t) \\ \cdots & \cdots \\ H_M: & r(t) = s_M(t) + n(t) \end{cases} \tag{107}$$

Corollary (LFs)

It can be proven that:

$$pdf_{r/H_i}(r(t)) = const \cdot \exp\left\{-\frac{1}{N_0}\int_0^{T_{cs}} (r(t) - s_i(t))^2 dt\right\} \tag{108}$$

$$= const \cdot \exp\left\{\begin{array}{l} -\frac{1}{N_0}\int_0^{T_{cs}} r(t)^2 dt \\ -\frac{1}{N_0}\int_0^{T_{cs}} s_i(t)^2 dt \\ +\frac{2}{N_0}\int_0^{T_{cs}} r(t) s_i(t) dt \end{array}\right\} \tag{109}$$

Optimum M-ary Architecture based on Decision Theory

It can be easily proven that:

$$D_i = \begin{cases} \arg\max_i \{ \underbrace{\text{pdf}_{r/H_i}(r(t))}_{=G_i(r)} \} & \text{ML} \\ \arg\max_i \{ \underbrace{\Pr(H_i) \times \text{pdf}_{r/H_i}(r(t))}_{=G_i(r)} \} & \text{MAP, or } \min(\text{pe}) \\ \arg\max_i \{ \underbrace{\text{weight}_i \times \text{pdf}_{r/H_i}(r(t))}_{=G_i(r)} \} & \text{minmax} \\ \arg\max_i \{ \underbrace{\text{weight}_i \times \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t))}_{=G_i(r)} \} & \text{Bayes} \end{cases} \quad (110)$$

$$\text{for } i = 1, \ldots, M$$

Based on Equ 109, the parameter $G_i(r)$ shown in the previous equation, can be simplified as follows:

$$\boxed{G_i(r) = \int_0^{T_{cs}} r(t) s_i^*(t) dt + \text{DC}_i}, \forall i, \ \forall \text{criterion} \quad (111)$$

- where DC_i $\forall i$, depends on the decision criterion.

- For all the criteria, DC_i is given as follows:

$$\mathrm{DC}_i \triangleq \begin{cases} -\frac{1}{2}E_i & \text{ML} \\ \frac{N_0}{2}\ln(\Pr(H_i)) - \frac{1}{2}E_i & \text{MAP, or } \underset{}{\text{min (pe)}} \\ \frac{N_0}{2}\ln(\text{weight}_i) - \frac{1}{2}E_i & \text{minmax} \\ \frac{N_0}{2}\ln(\text{weight}_i . \Pr(H_i)) - \frac{1}{2}E_i & \text{Bayes} \end{cases} \tag{112}$$

$$\text{with } E_i \triangleq \text{energy of } s_i(t) \tag{113}$$

- Based on 111 which is repeated below:

$$\boxed{G_i(r) = \int_0^{T_{cs}} r(t) s_i^*(t) dt + \mathrm{DC}_i}, \forall i, \forall \text{criterion},$$

Equ 110, can be implemented by the following optimum architecture:

OPTIMUM M-ary RECEIVER: CORREL. RECEIVER