Introduction to Array Processing

Olivier Besson

Institut Supérieur de l'Aéronautique et de l'Espace
SUPAERO
Contents

1 Introduction
   Multichannel processing
   A simple example: array processing with two antennas
   Array of sensors

2 Array processing model

3 Beamforming

4 Direction of arrival estimation
Context of multichannel processing

- Multichannel processing involves processing measurement vectors
\[ y(k) = [y_1(k) \ y_2(k) \ \ldots \ y_N(k)]^T \in \mathbb{C}^N \] which can be:

- \( N \) samples collected (possibly at the same time) on \( N \) different sensors, for instance the temperature measured in \( N \) different locations, the blood pressure of \( N \) different persons, etc.
- \( N \) samples taken on a single sensor over a time frame of \( NT_s \) where \( T_s \) is the sampling period. In radar \( N \) is the number of pulses sent by a radar each \( T_s \) seconds during a coherent processing interval.

- Analysis of such vectors should be understood as figuring out the relation between \( y_n(.) \) and \( y_m(.) \):
  - how correlated are the signals \( y_n(.) \) and \( y_m(.) \)?
  - can \( y_n(.) \) be explained by a subset of the variables?
**Context of array processing**

- In this course, we focus on signals received on an array of $N$ antennas placed at different locations with a view to enhance reception of signals coming from preferred directions.

- $y_n(k)$ represents the signal received at antenna number $n$ at time index number $k$.

- $n \in [1, N]$ is a spatial index as it is related to a particular position of an antenna in space and one is interested in analyzing or processing the signals $y_n(\cdot)$. 

---

**Introduction**

Multichannel processing
A generic multichannel problem

- Ubiquitous problem: the vectors $y(k)$ contain some information (usually buried in noise) one wants to retrieve.
- A classical multichannel problem is to detect/estimate a known signal $a$ that would be present in $y$ in addition to some noise $n$, i.e., decide whether $\alpha = 0$ or $\alpha \neq 0$ from observation of $y = \alpha a + n$.
- The simplest way is to design a linear filter $w$ such that

$$
\sum_{n=1}^{N} w_n^* y_n = \begin{bmatrix} w_1^* & w_2^* & \ldots & w_N^* \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = w^H y \simeq \alpha
$$

and possibly to test the energy $|w^H y|^2$. 

Introduction
Multichannel processing
Array processing with two antennas

- Consider two antennas receiving a signal $s(t)$ emitted by a source in the far-field

$$y_1(t) \approx As(t - t_0)$$
$$y_2(t) \approx As(t - t_0 - \Delta t)$$

- The time delay $\Delta t$ depends on the direction of arrival $\theta$ of $s(t)$ and on the relative (known) positions of the antennas:
  - if $\theta$ is known, one can obtain $s(t)$: spatial filtering (beamforming)
  - if one can estimate $\Delta t$ from $y_1(t)$ and $y_2(t)$, then $\theta$ follows: source localization.
Beamforming with two antennas

- For narrowband signals, a delay amounts to a phase shift. Hence

\[ y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = As(t) \begin{bmatrix} 1 \\ e^{i\phi} \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix} \]

- Let us use a linear filter to estimate \( As(t) \). The output is

\[ w_1^* y_1(t) + w_2^* y_2(t) = As(t)[w_1^* + w_2^* e^{i\phi}] + [w_1^* n_1(t) + w_2^* n_2(t)] \]

- A good filter maximizes the output signal to noise ratio (SNR) which, if \( n_1(t) \) and \( n_2(t) \) are uncorrelated and have same power, is given by

\[
\text{SNR} = \frac{|w_1^* + w_2^* e^{i\phi}|^2}{|w_1|^2 + |w_2|^2} \frac{|A|^2 P_s}{P_n}
\]

\[ \rightarrow \text{maximum for } w_2 = w_1 e^{i\phi} , \text{ so that } w^H y(t) \propto y_1(t) + y_2(t) e^{-i\phi}. \]
Array of sensors

Potentialities

Array of sensors offer an additional dimension (space) which enables one, possibly in conjunction with temporal or frequency filtering, to perform spatial filtering of signals:

1. source separation
2. direction finding

Fields of application

1. radar, sonar (detection, target localization, anti-jamming)
2. communications (system capacity improvement, enhanced signals reception, spatial focusing of transmissions, interference mitigation)
Contents

1 Introduction

2 Array processing model
   Principle
   Multichannel receiver
   Signals received on the array
   Covariance matrix
   Model limitations

3 Beamforming

4 Direction of arrival estimation
The array performs **spatial sampling** of a wavefront impinging from direction\((\theta, \phi)\).

Assumptions: homogeneous propagation medium, source in the far-field of the array \(\rightarrow\) plane wavefront.
Multichannel receiver

Array processing model

Multichannel receiver
Source signal (frequency domain)

\[
\tilde{X}(\omega) = S(\omega) \tilde{H}(\omega) - \omega_c \omega_c
\]
Signals and receiver

### Source signal (narrowband)

\[
\ddot{x}(t) = 2\text{Re}\left\{s(t)e^{i\omega_ct}\right\} \\
\triangleq \text{Re}\left\{\alpha(t)e^{i\phi(t)}e^{i\omega_ct}\right\} \\
= \alpha(t) \cos[\omega_c t + \phi(t)]
\]

\(\alpha(t)\) and \(\phi(t)\) stand for amplitude and phase of \(s(t)\), and have slow time-variations relative to \(f_c\).

### Channel response

Receive channel number \(n\) has impulse response \(\ddot{h}_n(t)\).
Model of received signals

- Signal received on $n$-th antenna

$$\tilde{y}_n(t) = \alpha \tilde{h}_n(t) * \tilde{x}(t - \tau_n) + \tilde{n}_n(t)$$

where $\tau_n$ is the propagation delay to $n$-th sensor.

- In frequency domain:

$$\tilde{Y}_n(\omega) = \alpha \tilde{H}_n(\omega) \tilde{X}(\omega) e^{-i\omega \tau_n} + \tilde{N}_n(\omega)$$

- After demodulation ($\omega \rightarrow \omega + \omega_c$) and lowpass filtering:

$$Y_n(\omega) = \alpha \tilde{H}_n(\omega + \omega_c) S(\omega) e^{-i(\omega + \omega_c) \tau_n} + \tilde{N}_n(\omega + \omega_c)$$

$$\approx \alpha \tilde{H}_n(\omega_c) S(\omega) e^{-i\omega_c \tau_n} + \tilde{N}_n(\omega + \omega_c)$$
Model of received signals

• Taking the inverse Fourier transform $\mathcal{F}^{-1}(Y_n(\omega))$ yields

$$y_n(t) \simeq \alpha \tilde{H}_n(\omega_c)s(t)e^{-i\omega_c\tau_n} + n_n(t)$$

• The snapshot writes

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_N(t) \end{bmatrix} = \alpha \begin{bmatrix} \tilde{H}_1(\omega_c)e^{-i\omega_c\tau_1} \\ \tilde{H}_2(\omega_c)e^{-i\omega_c\tau_2} \\ \vdots \\ \tilde{H}_N(\omega_c)e^{-i\omega_c\tau_N} \end{bmatrix} \begin{bmatrix} s(t) \\ n_1(t) \\ n_2(t) \\ \vdots \\ n_N(t) \end{bmatrix}$$

• Assuming all $\tilde{H}_n(\omega_c)$ are identical and absorbing $\alpha$ and $\tilde{H}_n(\omega_c)$ in $s(t)$, we simply write

$$\mathbf{y}(t) = \mathbf{a}(\theta)s(t) + \mathbf{n}(t)$$
Model of received signals

- The snapshot is then sampled (temporally) at rate $T_s$ to obtain the $N|K$ data matrix $\mathbf{Y} = \begin{bmatrix} \mathbf{y}(1) & \mathbf{y}(2) & \ldots & \mathbf{y}(K) \end{bmatrix}$:

- The $k$-th snapshot is given by

$$\mathbf{y}(k) = \mathbf{a}(\theta)s(k) + \mathbf{n}(k)$$

where $\mathbf{a}(\theta)$ is the vector of phase shifts, referred to as the steering vector since $\tau_n$ depends only on the directions(s) of arrival of the source.
Model of received signals

Snapshot at time index $k$:
The snapshot received in the presence of $P$ sources is given by

$$y(k) = \sum_{p=1}^{P} a(\theta_p) s_p(k) + n(k)$$

$$= \begin{bmatrix} a(\theta_1) & \ldots & a(\theta_P) \end{bmatrix} \begin{bmatrix} s_1(k) \\ \vdots \\ s_P(k) \end{bmatrix} + n(k)$$

$$= A(\theta)s(k) + n(k)$$
Steering vector

\[
\tau_n = \frac{1}{c} \left[ x_n \cos \theta \cos \phi + y_n \cos \theta \sin \phi + z_n \sin \theta \right]
\]

\[
a_n(\theta, \phi) = e^{i \frac{2\pi}{\lambda} \left[ x_n \cos \theta \cos \phi + y_n \cos \theta \sin \phi + z_n \sin \theta \right]}
\]
Uniform linear array (ULA)

Steering vector

\[ a(\theta) = \begin{bmatrix} 1 & e^{i2\pi f_s} & \ldots & e^{i2\pi(N-1)f_s} \end{bmatrix}^T ; \quad f_s = f_c \frac{d \sin \theta}{c} = \frac{d}{\lambda} \sin \theta \]

Shannon spatial sampling theorem

\[ |f_s| \leq 0.5 \Rightarrow d \leq \frac{\lambda}{2} \]
Covariance matrix

Definition
The covariance matrix is defined as

\[
\mathbf{R} = \mathbb{E} \{ \mathbf{y}(k) \mathbf{y}^H(k) \} = \mathbb{E} \left\{ \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_N(k) \end{bmatrix} \begin{bmatrix} y_1^*(k) & y_2^*(k) & \cdots & y_N^*(k) \end{bmatrix} \right\}
\]

Interpretation
The \((n, \ell)\) entry \(\mathbf{R}(n, \ell) = \mathbb{E} \{ y_n(k) y_\ell^*(k) \}\) measures the correlation between signals received at sensors \(n\) and \(\ell\), at the same time index \(k\).
Structure of the covariance matrix

Signals covariance matrix

The covariance matrix of the signal component is

$$ R = \mathcal{E} \{ A(\theta)s(k)s^H(k)A^H(\theta) \} = A(\theta)R_s A^H(\theta) $$

$$ = \sum_{p=1}^{P} P_p a(\theta_p) a^H(\theta_p) \quad (uncorrelated \ signals) $$

Provided that $R_s$ is full-rank (non coherent signals), the signal covariance matrix has rank $P$ and its range space is spanned by the steering vectors $a(\theta_p), \ p = 1, \ldots, P$.

Noise covariance matrix

Assuming spatially white noise (i.e., uncorrelated between channels) with same power on each channel, $\mathcal{E} \{ n(k)n^H(k) \} = \sigma^2 I$. 

Array processing model | Covariance matrix
Model limitations

\( y(k) = a(\theta)s(k) + n(k) \) is an idealized model of the signals received on the array. It does not account for:

- a possibly non homogeneous propagation medium which results in coherence loss and wavefront distortions. This leads to amplitude and phase variations along the array, i.e.
  \[ y_n(k) = g_n(k)e^{i\phi_n(k)}a_n(\theta)s(k) + n_n(k). \]
- uncalibrated arrays, i.e., different amplitude and phase responses for each channel.
- wideband signals for which a time delay does not amount to a simple phase shift. In the frequency domain, one has
  \[ y(f) = a_f(\theta)s(f) + n(f) \]
  with
  \[ a_f(\theta) = \begin{bmatrix} 1 & e^{-i2\pi f\tau(\theta)} & \ldots & e^{-i2\pi f(N-1)\tau(\theta)} \end{bmatrix}^T. \]
- possibly colored reception noise, i.e. \( \mathbb{E} \{ n(k)n^H(k) \} \neq \sigma^2 I. \)
Contents

1 Introduction

2 Array processing model

3 Beamforming
   - Principle
   - Array beampattern
   - Spatial filtering
   - Adaptive beamforming
   - Robust adaptive beamforming
   - Partially adaptive beamforming
   - Summary

4 Direction of arrival estimation
Spatial filtering

Principle: use a **linear combination of the sensors outputs** in order to point towards a looked direction.

\[
y_F(k) = \sum_{n=1}^{N} w^*_{n} y_n(k) \simeq a_s(k)
\]

Diagram: 
- \(s(t)\) and \(i(t)\) are signals.
- \(y_1(k), y_2(k), \ldots, y_N(k)\) are sensor outputs.
- \(w_1^*, w_2^*, \ldots, w_N^*\) are weights.
- The sum of the weighted sensor outputs results in \(y_F(k)\).

Beamforming

Principle
Array beampattern

- For any weight vector $w$, the corresponding array beampattern is defined as

$$G_w(\theta) = |g_w(\theta)|^2 = |w^H a(\theta)|^2$$

- For a uniform linear array, the natural beampattern, obtained as a simple sum ($w_n = 1$) of the sensors outputs, is given by

$$g(\theta) = \sum_{n=0}^{N-1} e^{i2\pi n \frac{d}{\lambda} \sin \theta} = e^{i\pi (N-1) \frac{d}{\lambda} \sin \theta} \frac{\sin \left[ \pi N \frac{d}{\lambda} \sin \theta \right]}{\sin \left[ \pi \frac{d}{\lambda} \sin \theta \right]}$$

$$G(\theta) = |g(\theta)|^2 = \left| \frac{\sin \left[ \pi N \frac{d}{\lambda} \sin \theta \right]}{\sin \left[ \pi \frac{d}{\lambda} \sin \theta \right]} \right|^2$$
ULA beampattern

Beampattern of the uniform linear array

\[ \theta_{3\text{dB}} \approx \frac{0.9\lambda}{N d} \]
Windowing

ULA beampattern with windowing

- Rect
- Cheb 30dB
- Cheb 50dB

Angle of arrival

Beamforming

Array beampattern
Beamforming

Objective

We aim at pointing towards a given direction in order to enhance reception of the signals impinging from this direction, and to possibly mitigate interference located at other directions.

Principle

Each sensor output is weighted by $w_n^*$ before summation:

$$y_F(k) = \sum_{n=1}^{N} w_n^* y_n(k) = \begin{bmatrix} w_1^* & w_2^* & \cdots & w_N^* \end{bmatrix} y(k) = \mathbf{w}^H y(k).$$

Question

How to choose $\mathbf{w}$ such that, if $y(k) = a(\theta_s)s(k) + \cdots$ then at the output $y_F(k) \simeq \alpha s(k)$?
Conventional beamforming

Conventional beamforming: $w \propto a(\theta_s)$

$$y_F(k) = a^H(\theta_s)a(\theta_s)s(k) \quad [w = a(\theta_s), \text{1 source at } \theta_s]$$

$$= \sum_{n=0}^{N-1} e^{-i2\pi \frac{d}{\lambda} n \sin \theta_s} \times e^{+i2\pi \frac{d}{\lambda} n \sin \theta_s} s(k)$$

$$= \sum_{n=0}^{N-1} s(k) = Ns(k)$$

so that the gain towards $\theta_s$ is **maximal** and equal to $N$. The beamformer $w_{\text{CBF}} = a(\theta_s)/[a^H(\theta_s)a(\theta_s)]$ is referred to as the conventional beamformer.

**Principle**

One compensates for the phase shift induced by propagation from direction $\theta_s$ and then sum **coherently**.
Array beampattern with conventional beamforming
**SNR improvement**

**Before beamforming**

\[ y(k) = a_s s(k) + n(k); \quad \text{SNR}_{\text{elem}} \triangleq \frac{\mathcal{E} \left\{ \left| s(k) \right|^2 \right\}}{\mathcal{E} \left\{ \left| n(k) \right|^2 \right\}} = \frac{P}{\sigma^2}. \]

**After beamforming**

\[ y_F(k) = w^H y(k) = w^H a_s s(k) + w^H n(k) \]

\[ \text{SNR}_{\text{array}} = \frac{|w^H a_s|^2}{\|w\|^2} \text{SNR}_{\text{elem}} \leq \|a_s\|^2 \text{SNR}_{\text{elem}} = N \times \text{SNR}_{\text{elem}} \]

*with equality if* \( w \propto a_s \).

**White noise array gain**

For any \( w \) such that \( w^H a_s = 1 \), the **white noise array gain** is

\[ A_{\text{WN}} = \frac{\text{SNR}_{\text{array}}}{\text{SNR}_{\text{elem}}} = \|w\|^{-2} \leq N. \]
Conventional beamforming versus adaptive beamforming

Conventional beamforming

The conventional beamformer is optimal in white noise: it amounts to minimize $w^H w$ (the output power in white noise) under the constraint $w^H a(\theta_s) = 1$. Any other direction is deemed to be equivalent $\Rightarrow$ it does not take into account other signals (interference) present in some directions.

Adaptive beamforming

Adaptive beamforming takes into account these other signals. It consists in minimizing the output power $E \left\{ \left| w^H y(k) \right|^2 \right\}$ while maintaining a unit gain towards looked direction $\Rightarrow$ tends to place nulls towards interfering signals.
Adaptive beamforming

Beamforming-filtering in the presence of interference

• The received (input) signal in the presence of interference and noise is given by

\[ y(k) = a_s s(k) + y_I(k) + n(k) \]

where \( a_s \) is the actual SOI steering vector.

• At the output of the beamformer

\[ y(k) = a_s s(k) + y_I(k) + n(k) \]

\[ w^H a_s s(k) + w^H [y_I(k) + n(k)] \]
Signal to interference plus noise ratio (SINR)

Definition of SINR

For a given beamformer $\mathbf{w}$, the usual figure of merit is the signal to interference plus noise ratio (SINR), defined as

$$
\text{SINR}(\mathbf{w}) = \frac{\mathcal{E} \left\{ |\mathbf{w}^H \mathbf{a}_s s(k)|^2 \right\}}{\mathcal{E} \left\{ |\mathbf{w}^H [\mathbf{y}_I(k) + \mathbf{n}(k)]|^2 \right\}}
$$

$$
= \frac{P_s |\mathbf{w}^H \mathbf{a}_s|^2}{\mathbf{w}^H \mathbf{C} \mathbf{w}}
$$

where $\mathbf{C} = \mathcal{E} \left\{ [\mathbf{y}_I(k) + \mathbf{n}(k)] [\mathbf{y}_I(k) + \mathbf{n}(k)]^H \right\}$ stands for the interference plus noise covariance matrix.
Optimal beamformer: SINR maximization

Optimal beamformer

Maximize SINR while ensuring a unit gain towards $a_s$:

$$\min_{w} w^H C w \text{ subject to } w^H a_s = 1$$

($\text{optimal}$)

$$w_{opt} = \frac{C^{-1} a_s}{a_s^H C^{-1} a_s} \rightarrow SINR_{opt} = P_s a_s^H C^{-1} a_s$$

Remarks

• Principle is to minimize output power (when input $= y_I + n$) under the constraint that the actual steering vector $a_s$ goes non distorted.

• Neither $a_s$ nor $C$ will be known in practice: the actual steering vector may be different from its expected value and $C$ needs to be estimated from data (which contain $y_I + n$).
Minimum Variance Distortionless Response (MVDR)

**Principle**

Minimize output power \((\text{when input } = y_I + n)\) under the constraint that the **assumed** steering vector goes non distorted.

**Minimization problem and solution**

\[
\min_{\mathbf{w}} \mathbf{w}^H \mathbf{C} \mathbf{w} \quad \text{subject to} \quad \mathbf{w}^H \mathbf{a}_0 = 1 \\
\text{(MVDR)}
\]

where \(\mathbf{a}_0\) is the **assumed** steering vector of the signal of interest (SoI). The solution is given by

\[
\mathbf{w}_{\text{MVDR}} = \frac{\mathbf{C}^{-1} \mathbf{a}_0}{\mathbf{a}_0^H \mathbf{C}^{-1} \mathbf{a}_0}
\]
Minimum Power Distortionless Response (MPDR)

**Principle**

Minimize output power (when input = $a_s s + y_I + n$) under the constraint that the assumed steering vector goes non distorted:

$$\min_w w^H R w \text{ subject to } w^H a_0 = 1 \quad \text{(MPDR)}$$

where $R(= C + P_s a_s a_s^H)$ stands for the **signal plus interference plus noise** covariance matrix.

**Solution**

$$w_{MPDR} = \frac{R^{-1} a_0}{a_0^H R^{-1} a_0}$$
# Summary of adaptive beamformers (known covariance matrices)

<table>
<thead>
<tr>
<th>Beamformer</th>
<th>Principle</th>
<th>Weight vector</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Optimal</strong></td>
<td>( \min_w w^H C w \text{ s.t. } w^H a_s = 1 )</td>
<td>( w_{\text{opt}} = \frac{C^{-1}a_s}{a_s^H C^{-1}a_s} )</td>
</tr>
<tr>
<td><strong>MVDR</strong></td>
<td>( \min_w w^H C w \text{ s.t. } w^H a_0 = 1 )</td>
<td>( w_{\text{MVDR}} = \frac{C^{-1}a_0}{a_0^H C^{-1}a_0} )</td>
</tr>
<tr>
<td><strong>MPDR</strong></td>
<td>( \min_w w^H R w \text{ s.t. } w^H a_0 = 1 )</td>
<td>( w_{\text{MPDR}} = \frac{R^{-1}a_0}{a_0^H R^{-1}a_0} )</td>
</tr>
</tbody>
</table>

- \( a_s (a_0) \) the actual (assumed) steering vector
- \( C = \text{cov}(y_I + n) \) and \( R = \text{cov}(a_s s + y_I + n) \)
CBF and optimal (MVDR) beampatterns
CBF vs MVDR: the case of a single interference

**Derivation of SINR**

In the case $\mathbf{C} = P_j \mathbf{a}_j \mathbf{a}_j^H + \sigma^2 \mathbf{I}$ with $\text{INR} = \frac{P_j}{\sigma^2} \gg 1$, it can be shown that

\[
\text{SINR}_{\text{CBF}} \simeq \frac{P_s}{\sigma^2} \times \frac{1}{g \times \text{INR}}; \quad \text{SINR}_{\text{opt}} \simeq \frac{P_s}{\sigma^2} \times N(1 - g)
\]

with $g = \cos^2 (\mathbf{a}_s, \mathbf{a}_j) = |\mathbf{a}_s^H \mathbf{a}_j|^2 / (\mathbf{a}_s^H \mathbf{a}_s)(\mathbf{a}_j^H \mathbf{a}_j)$.

**Remarks**

- With CBF, the SINR decreases when $P_j$ increases while it is independent of $P_j$ with adaptive beamforming.
- The SINR decreases when $\mathbf{a}_j \rightarrow \mathbf{a}_s$ ($g \rightarrow 1$).
CBF vs MVDR

- From Kantorovich’s inequality

\[
1 \leq \frac{\text{SINR}(\mathbf{w}_{\text{opt}})}{\text{SINR}(\mathbf{w}_{\text{CBF}})} = \frac{(\mathbf{a}_s^H \mathbf{C}^{-1} \mathbf{a}_s)(\mathbf{a}_s^H \mathbf{C} \mathbf{a}_s)}{(\mathbf{a}_s^H \mathbf{a}_s)^2} \leq \frac{(\lambda_{\text{min}}(\mathbf{C}) + \lambda_{\text{max}}(\mathbf{C}))^2}{4 \lambda_{\text{min}}(\mathbf{C}) \lambda_{\text{max}}(\mathbf{C})}
\]

- Adaptive beamforming is adequate if $\frac{\lambda_{\text{max}}(\mathbf{C})}{\lambda_{\text{min}}(\mathbf{C})} \gg 1$. 
Generalized Sidelobe Canceler

- The GSC structure can be represented as

\[
\begin{align*}
\mathbf{y}(k) &= a_0 s(k) + n(k) \quad \mathbf{a}_0 \\
\mathbf{i}(k) &= a_0 s(k) + n(k) \quad \mathbf{a}_0 \\
\mathbf{n}(k) &= a_0 s(k) + n(k) \quad \mathbf{a}_0 \\
\mathbf{d}(k) &= s(k) + i_1(k) + n_1(k) \quad \mathbf{d}(k) \quad \mathbf{s}(k) \quad \mathbf{i}_1(k) \quad \mathbf{n}_1(k) \\
\mathbf{w_a} &= \mathbf{w_a} \quad \mathbf{w_a} \\
\mathbf{B} &= \mathbf{B} \\
\mathbf{z}(k) &= i_2(k) + n_2(k) \quad \mathbf{i}_2(k) \quad \mathbf{n}_2(k) \\
\mathbf{s}(t) &= \mathbf{s}(t) \\
\mathbf{i}_2(t) &= \mathbf{i}_2(t) \\
\mathbf{z}(k) &= \mathbf{z}(k) \\
\mathbf{d}(k) - \mathbf{w_a}^H \mathbf{z}(k) &= \mathbf{d}(k) - \mathbf{w_a}^H \mathbf{z}(k)
\end{align*}
\]

where the \((N - 1)\) columns of \(\mathbf{B}\) form a basis of the subspace orthogonal to \(\mathbf{a}_0\), i.e., \(\mathbf{B}^H \mathbf{a}_0 = 0\).

- The \((N - 1)\) auxiliary channels \(\mathbf{z}(k)\) are free of signal and enable one to infer the part of interference that went through the CBF.

- \(\mathbf{w_a}\) enables one to estimate, from \(\mathbf{z}(k)\), the part of interference \(\mathbf{i}_1(k)\) contained in \(\mathbf{d}(k)\) since \(\mathbf{i}_1(k)\) is correlated with \(\mathbf{z}(k)\) through \(\mathbf{i}_2(k)\).
Generalized Sidelobe Canceler

• The GSC structure decomposes $w$ into a component along $a_0$ and a component orthogonal to $a_0$, i.e., $w = \alpha a_0 - w_\perp$:

$\begin{align*}
y(k) &\rightarrow \frac{a_0}{a_0^H a_0} \rightarrow \frac{d(k)}{s(k)} \rightarrow \left\lfloor \begin{array}{c}
a_0 \\vline \\i_1(k) \\
N \vline N - 1
\end{array} \right. \\
&\rightarrow B \rightarrow \frac{z(k)}{i_2(k)} \rightarrow \left\lfloor \begin{array}{c}
N - 1 \vline 1
\end{array} \right. \\
&\rightarrow w_a \rightarrow d(k) - w_\perp^H z(k)
\end{align*}$

• The component along $a_0$ ensures that the constraint is fulfilled since

$$w^H a_0 = \alpha^* a_0^H a_0 - w_\perp^H a_0 = \alpha^* a_0^H a_0 + 0 \Rightarrow \alpha = (a_0^H a_0)^{-1}$$

• The orthogonal component $w_\perp = B w_a$ is chosen to minimize output power, in an unconstrained way.
Generalized Sidelobe Canceler

• Minimization of the output power can be achieved by solving one of the two following equivalent problems:

\[
\begin{align*}
\min_{w} & \quad w^H C w \\
\text{subject to} & \quad w^H a_0 = 1 \\
\min_{w_a} & \quad (w_{\text{CBF}} - B w_a)^H C (w_{\text{CBF}} - B w_a)
\end{align*}
\]

direct form, constrained \quad GSC form, unconstrained

• The MVDR beamformer in its GSC form is given by

\[
w_{\text{GSC}} = w_{\text{CBF}} - B w_a^*
\]

where \(w_a^*\) solves the above minimization problem.
Generalized Sidelobe Canceler

**Derivation of $w_a^*$**

- The power at the output of the beamformer is given by

$$
E \left\{ \left| d(k) - w_a^H z(k) \right|^2 \right\} = E \left\{ |d(k)|^2 \right\} - w_a^H r_{dz} - r_{dz}^H w_a + w_a^H R_z w_a
$$

$$
= \left[ w_a - R_z^{-1} r_{dz} \right]^H R_z \left[ w_a - R_z^{-1} r_{dz} \right]
$$

$$
+ E \left\{ |d(k)|^2 \right\} - r_{dz}^H R_z^{-1} r_{dz}
$$

with $r_{dz} = E \left\{ z(k) d^*(k) \right\}$ and $R_z = E \left\{ z(k) z(k)^H \right\}$.

- The weight vector which minimizes output power is thus

$$
w_a^* = R_z^{-1} r_{dz}
$$
Generalized Sidelobe Canceler

• The GSC form of the weight vector is given by

\[ w_{\text{GSC}} = w_{\text{CBF}} - BR_z^{-1}r_dz \]
\[ = w_{\text{CBF}} - B (B^H R_y B)^{-1} B^H R_y w_{\text{CBF}} \]  \hspace{1cm} \text{(GSC)}

where \( R_y = R \) in a MPDR scenario and \( R_y = C \) in a MVDR scenario.

• Since they solve the \textit{same problem} \( w_{\text{GSC}} = (a_0^H R^{-1} a_0)^{-1} R^{-1} a_0 \).

• The SINR is inversely proportional to the output power when \( R_y = C \), i.e.,

\[ \text{SINR}_{\text{GSC}} = P_s \left[ w_{\text{CBF}}^H C w_{\text{CBF}} - r_dz^H R_z^{-1} r_dz \right]^{-1} \]
The optimal, MVDR and MPDR beamformers are equivalent if and only if

$$\min_{\mathbf{w}} \mathbf{w}^H \left( \mathbf{C} + P_s \mathbf{a}_s \mathbf{a}_s^H \right) \mathbf{w} \text{ subject to } \mathbf{w}^H \mathbf{a}_0 = 1 \quad \text{(MPDR)}$$

$$\equiv \min_{\mathbf{w}} \mathbf{w}^H \mathbf{C} \mathbf{w} \text{ subject to } \mathbf{w}^H \mathbf{a}_0 = 1 \quad \text{(MVDR)}$$

$$\equiv \min_{\mathbf{w}} \mathbf{w}^H \mathbf{C} \mathbf{w} \text{ subject to } \mathbf{w}^H \mathbf{a}_s = 1 \quad \text{(opt)}$$

which is true only when the 2 following conditions are satisfied:

1. the assumed steering vector $\mathbf{a}_0$ coincides with the actual steering vector $\mathbf{a}_s$: in practice, uncalibrated arrays or a pointing error lead to $\mathbf{a}_0 \neq \mathbf{a}_s$;

2. the covariance matrix $\mathbf{R}$ is known: in practice, one needs to estimate it which results in estimation errors $\hat{\mathbf{R}} - \mathbf{R}$.

$$\Rightarrow$$ It ensues that degradation compared to $\text{SINR}_{\text{opt}}$ is unavoidable in practice, and it can be quite different between MPDR and MVDR.
Influence of a steering vector error (MVDR)

- We assume that the SoI steering vector is $a_0$ while it is actually $a_s$.
- The SINR obtained with $w_{\text{MVDR}} = \left(a_0^H C^{-1} a_0\right)^{-1} C^{-1} a_0$ becomes

$$
\text{SINR}_{\text{MVDR}} = \frac{P_s \left| w_{\text{MVDR}}^H a_s \right|^2}{w_{\text{MVDR}}^H C w_{\text{MVDR}}} = P_s \frac{\left| a_0^H C^{-1} a_s \right|^2}{a_0^H C^{-1} a_0}
$$

$$
= \text{SINR}_{\text{opt}} \times \frac{\left| a_0^H C^{-1} a_s \right|^2}{(a_0^H C^{-1} a_0)(a_s^H C^{-1} a_s)}
$$

$$
= \text{SINR}_{\text{opt}} \times \cos^2 \left(a_s, a_0; C^{-1}\right)
$$

$$
\leq \text{SINR}_{\text{opt}}
$$
Influence of a steering vector error (MPDR)

- The MPDR beamformer can be written as
  \[ w_{\text{MPDR}} = \frac{\mathbf{R}^{-1}\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{R}^{-1} \mathbf{a}_0}; \quad \mathbf{R} = P_s \mathbf{a}_s \mathbf{a}_s^H + \mathbf{C} \]

- Its SINR is decreased compared to that of the MVDR, viz
  \[ \text{SINR}_{\text{MPDR}} = \frac{\text{SINR}_{\text{MVDR}}}{1 + (2\text{SINR}_{\text{opt}} + \text{SINR}_{\text{opt}}^2) \sin^2 (\mathbf{a}_s, \mathbf{a}_0; \mathbf{C}^{-1})} \leq \text{SINR}_{\text{MVDR}}. \]

- The degradation is more important as SINR_{\text{opt}} (hence $P_s$) increases.
Influence of a steering vector error on beampatterns

\[ \theta_0 - \theta_s = 2^\circ \]
Influence of a steering vector error on SINR and WNAG

SINR loss when $a_0 \neq a_s$

White noise array gain when $a_0 \neq a_s$
Case of an uncalibrated array

- Let us consider an uncalibrated array with actual steering vector

\[
\tilde{a}_n(\theta) = (1 + g_n)e^{i\phi_n}a_n(\theta)
\]

where \(\{g_n\}\) and \(\{\phi_n\}\) are independent random gains and phases.

- For any beamformer \(w\), the average value of the resulting beampattern \(\tilde{G}_w(\theta) = |w^H\tilde{a}(\theta)|^2\) is related to the nominal beampattern \(G_w(\theta) = |w^Ha(\theta)|^2\) through

\[
E\left\{\tilde{G}_w(\theta)\right\} = |\gamma|^2 G_w(\theta) + \left[1 + \sigma_g^2 - |\gamma|^2\right] \|w\|^2
\]

where \(\sigma_g^2 = E\left\{|g_n|^2\right\}\) and \(\gamma = E\left\{e^{i\phi_n}\right\}\).

- The term proportional to \(\|w\|^2\) leads to sidelobe level increase \(\Rightarrow\) better to have high white noise array gain (low \(\|w\|^2\)).
Influence of a finite number of snapshots

• In practice, $K$ snapshots are available:

\[ y(k) = a_s s(k) + y_I(k) + n(k); \quad k = 1, \ldots, K \]

• The covariance matrices are thus estimated and subsequently one can compute the corresponding beamformers as

\[
\hat{R} = \frac{1}{K} \sum_{k=1}^{K} y(k)y^H(k) \quad \rightarrow \quad w_{\text{MPDR}}^{\text{smi}} = \frac{\hat{R}^{-1}a_0}{a_0^H \hat{R}^{-1}a_0}
\]

\[
\hat{C} = \frac{1}{K} \sum_{k=1}^{K} y_{i+n}(k)y_{i+n}^H(k) \quad \rightarrow \quad w_{\text{MVDR}}^{\text{smi}} = \frac{\hat{C}^{-1}a_0}{a_0^H \hat{C}^{-1}a_0}
\]

where $^{\text{smi}}$ stands for “sample matrix inversion”.

Beamforming
Adaptive beamforming
Influence of a finite number of snapshots

- The sample beamformers $\mathbf{w}_{\text{M-DR}}^{\text{smi}}$ will differ from their ensemble counterparts $\mathbf{w}_{\text{M-DR}}$ since $\hat{\mathbf{R}} = \mathbf{R} + \Delta \mathbf{R}$ and $\hat{\mathbf{C}} = \mathbf{C} + \Delta \mathbf{C}$.

- The weight vectors $\mathbf{w}_{\text{M-DR}}^{\text{smi}}$ are random and so are their corresponding signal to noise ratios

  $$\text{SINR} (\mathbf{w}_{\text{MPDR}}^{\text{smi}}) = P_s \frac{|a_0^H \hat{\mathbf{R}}^{-1} a_s|^2}{a_0^H \hat{\mathbf{R}}^{-1} \mathbf{C} \hat{\mathbf{R}}^{-1} a_0}$$

  $$\text{SINR} (\mathbf{w}_{\text{MVDR}}^{\text{smi}}) = P_s \frac{|a_0^H \hat{\mathbf{C}}^{-1} a_s|^2}{a_0^H \hat{\mathbf{C}}^{-1} \mathbf{C} \hat{\mathbf{C}}^{-1} a_0}$$

- Important issue is speed of convergence, i.e., how large should $K$ be for $\text{SINR} (\mathbf{w}_{\text{MPDR}}^{\text{smi}})$ or $\text{SINR} (\mathbf{w}_{\text{MVDR}}^{\text{smi}})$ to be “close” to $\text{SINR}_{\text{opt}}$?
SINR loss with finite number of snapshots (MVDR)

- When \( a_0 = a_s \), the \textbf{SINR loss} \( \rho_{\text{MVDR}} \in [0, 1] \)

\[
\rho_{\text{MVDR}} = \frac{\text{SINR} (w_{\text{SMI-MVDR}})}{\text{SINR} (w_{\text{opt}})} \xrightarrow{d} \left[ 1 + \frac{\chi^2_{2(N-1)}(0)}{\chi^2_{2(K-N+2)}(0)} \right]^{-1}
\]

follows a \textbf{beta distribution}, i.e.,

\[
p(\rho_{\text{MVDR}}) = \frac{\Gamma(K + 1)}{\Gamma(K - N + 2)\Gamma(N - 1)} \rho_{\text{MVDR}}^{K-N+1} (1 - \rho_{\text{MVDR}})^{N-2}
\]

- The expected value is \( \mathcal{E} \{ \rho_{\text{MVDR}} \} = (K - N + 2)/(K + 1) \), so that \( \text{SINR} (w_{\text{SMI-MVDR}}) \) is (on average) within 3dB of the optimal SINR for \( K_{\text{MVDR}} = 2N - 3 \).
SINR loss with finite number of snapshots (MVDR)

![Distribution of MVDR SINR loss]

- Green dashed line: $K = N$
- Blue dotted line: $K = 3N/2$
- Red dashed line: $K = 2N$
- Black line: $K = 4N$

Value of SINR loss (decibels)

Probability density function
SINR loss with finite number of snapshots (MPDR)

- As for $\rho_{\text{MPDR}}$ one has

  $$\rho_{\text{MPDR}} = \frac{\text{SINR} (w_{\text{MPDR}}^{\text{smi}})}{\text{SINR} (w_{\text{opt}})} = d \left[ 1 + (1 + \text{SINR}_{\text{opt}}) \frac{\chi^2_{2(N-1)}(0)}{\chi^2_{2(K-N+2)}(0)} \right]^{-1}$$

- The distribution of $\rho_{\text{MPDR}}$ is

  $$p(\rho_{\text{MPDR}}) = \frac{\Gamma(K + 1)(1 + \text{SINR}_{\text{opt}})^{K-N+2}}{\Gamma(N - 1)\Gamma(K - N + 2)} \frac{\rho_{\text{MPDR}}^{K-N+1}(1 - \rho_{\text{MPDR}})^{N-2}}{(1 + \rho_{\text{MPDR}}\text{SINR}_{\text{opt}})^{K+1}}$$

- The average number of snapshots to achieve the optimal SINR within 3dB is about

  $$K_{\text{MPDR}} \simeq (N - 1) [1 + \text{SINR}_{\text{opt}}]$$

  where $\text{SINR}_{\text{opt}} \simeq N \left( \frac{P_s}{\sigma^2} \right)$. In general, $K_{\text{MPDR}} \gg K_{\text{MVDR}}$. 

Beamforming  Adaptive beamforming
Beampatterns with finite number of snapshots

![Beampatterns CBF-MPDR-MVDR](image)

- **MVDR**
- **MPDR**
- **CBF**

Beampatterns CBF-MPDR-MVDR

- **dB**

**Angle of arrival**

N=10, K=20

**N=10, K=20**

Beamforming

Adaptive beamforming
Distribution of SINR loss

![Graph showing distribution of MPDR SINR loss]

- Red dotted line: MVDR
- Blue solid line: MPDR $SINR_{opt} = -10$ dB
- Green dashed line: MPDR $SINR_{opt} = 0$ dB
- Cyan dash-dotted line: MPDR $SINR_{opt} = 10$ dB

$N = 10, K = 20$
How to make MPDR more robust?

Observations

- Estimation of covariance matrices leads to a significant SINR loss (especially for the MPDR beamformer) due to
  - the interference being less eliminated
  - a sidelobe level increase which results in a lower white noise gain.
- In case of uncalibrated arrays, steering vector errors are all the more emphasized that the white noise gain is low (or $\|w\|^2$ large).
How to make MPDR more robust?

White noise array gain and SINR

A possible remedy

Enforce a minimal white noise array gain or equivalently restrain $\|w\|^2$ in order to make the MPDR beamformer more robust.
Diagonal loading

Principle

One tries to solve

$$\min_w w^H \hat{R} w \text{ subject to } w^H a_0 = 1 \text{ and } \|w\|^2 = A_{WN}^{-1}(\geq N^{-1})$$

Solution

The Lagrangian is given by (with $\lambda \in \mathbb{C}$ and $\mu \in \mathbb{R}$)

$$L(w, \lambda, \mu) = w^H \hat{R} w + \lambda (w^H a_0 - 1) + \lambda^* (a_0^H w - 1) + \mu (w^H w - A_{WN}^{-1})$$

$$= \left[ w + \lambda \left( \hat{R} + \mu I \right)^{-1} a_0 \right]^H \left( \hat{R} + \mu I \right) \left[ w + \lambda \left( \hat{R} + \mu I \right)^{-1} a_0 \right]$$

$$- \lambda - \lambda^* - \mu A_{WN}^{-1} - |\lambda|^2 a_0^H \left( \hat{R} + \mu I \right)^{-1} a_0.$$
Diagonal loading

Solution

The solution thus takes the form \( \mathbf{w}_{\text{MPDR-DL}} = -\lambda \left( \hat{\mathbf{R}} + \mu \mathbf{I} \right)^{-1} \mathbf{a}_0 \). Since \( \mathbf{w}_{\text{MPDR-DL}}^H \mathbf{a}_0 = 1 \), it follows that

\[
\mathbf{w}_{\text{MPDR-DL}} = \frac{(\hat{\mathbf{R}} + \mu \mathbf{I})^{-1} \mathbf{a}_0}{\mathbf{a}_0^H (\hat{\mathbf{R}} + \mu \mathbf{I})^{-1} \mathbf{a}_0}
\]

and \( \mu \) is selected such that \( \| \mathbf{w}_{\text{MPDR-DL}} \|^{-2} = A_{WN} \).
Diagonal loading: adaptivity versus robustness

 Beamforming

Robust adaptive beamforming
Choice of loading level

Many different possibilities have been proposed to set the loading level:

- set $A_{\text{WN}}$ (slightly below $N$) and compute $\mu$ from $\|w_{\text{MPDR-DL}}\|^{-2} = A_{\text{WN}}$.
- set $\mu$ directly, generally a few decibels above white noise level (see discussion next slide about beampatterns and eigenvalues).
- set $\mu$ using the theory of ridge regression, which enables one to compute $\mu$ from data.
- use that diagonal loading is the solution to the following problem
  \[
  \max_{P, \mathbf{a}} \hat{\mathbf{R}} - P\mathbf{a}\mathbf{a}^H \quad \text{for} \quad \|\mathbf{a} - \mathbf{a}_0\|^2 \leq \varepsilon^2
  \]
  and compute $\mu$ from $\varepsilon$.
- set $A_{\text{WN}}$ and compute directly the diagonally loaded beamformer in GSC form without necessarily computing $\mu$.
- ...
An interpretation of diagonal loading and the choice of $\mu$

- The array beampattern with the **true** covariance matrix is given by

$$g(\theta) = \frac{\alpha}{\sigma^2} \left\{ a_0^H a(\theta) - \sum_{n=1}^{J} \frac{\lambda_n}{\lambda_n + \sigma^2} \left[ a_0^H u_n \right] u_n^H a(\theta) \right\}$$

- The array beampattern with an **estimated** covariance matrix becomes

$$g_{\text{smi}}(\theta) = \frac{\alpha}{\hat{\lambda}_{\text{min}}} \left\{ a_0^H a(\theta) - \sum_{n=1}^{N} \frac{\hat{\lambda}_n}{\hat{\lambda}_n + \hat{\lambda}_{\text{min}}} \left[ a_0^H \hat{u}_n \right] \hat{u}_n^H a(\theta) \right\}$$

- Degradation is due to $\hat{\lambda}_{J+1} \neq \hat{\lambda}_{J+2} \neq \cdots \hat{\lambda}_N = \hat{\lambda}_{\text{min}}$.

- Replacing $\hat{R}$ by $\hat{R} + \mu I$ enables one to equalize the eigenvalues, provided that $\mu \gg \sigma^2$ and $\mu < \lambda_J$. 

---

**Beamforming**

**Robust adaptive beamforming**
Diagonal loading: SINR versus number of snapshots

SINR versus K – Diagonal loading

- optimum
- MPDR
- MPDR−DL=5dB
- MPDR−DL=10dB

Number of snapshots vs. SINR for different loading methods.
Diagonal loading: beampatterns

Beampattern of MPDR with diagonal loading

$N=10, K=20$

- $\text{MPDR}$
- $\text{MPDR-} DL=5\, \text{dB}$
- $\text{MPDR-} DL=10\, \text{dB}$

Angle of arrival (°)

$\text{dB}$
Influence of the loading level on SINR and WNAG

![SINR vs. Diagonal Loading Level](image1)

![White Noise Gain vs. Diagonal Loading Level](image2)

Beamforming

Robust adaptive beamforming
Linearly constrained beamforming

- To mitigate pointing errors, one can resort to multiple constraints, i.e. solve the problem

\[
\min w^H C w \text{ subject to } Z^H w = d
\]

whose solution is \( w = C^{-1}Z \left( Z^H C^{-1}Z \right)^{-1} d \).

- One can use a unit gain constraint around the presumed DOA or a smoothness constraint:

\[
Z = \begin{bmatrix} a(\theta_0) & a(\theta_0 + \delta_1) & \cdots & a(\theta_0 + \delta_L) \end{bmatrix}, \quad d = [1 \ 1 \ \cdots \ 1]^T
\]

\[
Z = \begin{bmatrix} a(\theta_0) & \frac{\partial a(\theta)}{\partial \theta} \bigg|_{\theta_0} & \cdots & \frac{\partial^L a(\theta)}{\partial \theta^L} \bigg|_{\theta_0} \end{bmatrix}, \quad d = [1 \ 0 \ \cdots \ 0]^T
\]
Partially adaptive beamforming

Principle
Perform beamforming in a lower dimensional subspace.

Structures
- Direct form:

\[
\begin{align*}
y(k) & \quad \rightarrow \quad T \quad \rightarrow \quad \tilde{y}(k) \quad \rightarrow \quad \tilde{w} \quad \rightarrow \quad \tilde{w}^H \tilde{y}(k)
\end{align*}
\]

- GSC form \((T = [\mathbf{w}_{\text{CBF}} \quad \mathbf{BU}])\):

\[
\begin{align*}
y(k) & \quad \rightarrow \quad \mathbf{w}_{\text{CBF}} \quad \rightarrow \quad d(k) \quad \rightarrow \quad \tilde{y}(k) \quad \rightarrow \quad \tilde{w}_{\alpha} \quad \rightarrow \quad d(k) - \tilde{w}_{\alpha}^H \tilde{z}(k)
\end{align*}
\]
Motivation for partially adaptive beamforming

The optimal beamformer when $C = \sum_{j=1}^{J} P_j a_j a_j^H + \sigma^2 I$

- With this low-rank + scaled identity matrix form, one has

$$C = \sum_{n=1}^{J} (\lambda_n + \sigma^2) u_n u_n^H + \sigma^2 \sum_{n=J+1}^{N} u_n u_n^H$$

where $\mathcal{R}\{u_1, \cdots, u_J\} = \mathcal{R}\{a_1, \cdots, a_J\}$.

- The optimal beamformer $w_{\text{opt}} \propto C^{-1} a_s$ can then be rewritten as

$$w_{\text{opt}} \propto w_{\text{CBF}} - \sum_{n=1}^{J} \frac{\lambda_n}{\lambda_n + \sigma^2} \left[u_n^H w_{\text{CBF}}\right] u_n$$
Motivation for partially adaptive beamforming

Interpretation of $\mathbf{w}_{opt}$ when $\mathbf{C} = \sum_{j=1}^{J} P_j \mathbf{a}_j \mathbf{a}_j^H + \sigma^2 \mathbf{I}$

- The optimal beamformer amounts to subtract from the CBF a linear combination of the $J$ principal eigenvectors of $\mathbf{C}$:

  \[
  y(k) \xrightarrow{\mathbf{w}_{CBF}} \left[ \begin{array}{c} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_J \end{array} \right] \xrightarrow{N|J} \left[ \begin{array}{c} \mathbf{u}_1^H \mathbf{w}_{CBF} \\ \vdots \\ \mathbf{u}_J^H \mathbf{w}_{CBF} \end{array} \right] \xrightarrow{J|1} \alpha \mathbf{w}_{opt}^H y(k)
  \]

- **The optimal beamformer is a partially adaptive beamformer.**

- The “eigenbeams” are steered towards interference in order to suppress them from the output of the conventional beamformer.
Beampatterns (CBF and eigenvectors)
Beampatterns (CBF and eigenvectors)

- CBF
- 1st eigen-beam
- 2nd eigen-beam
- MVDR

Beamforming
Partially adaptive beamforming
Expression of the partially adaptive beamformer

**Direct form**

- New snapshots after transformation $\tilde{y}(k) = T^H y(k)$ whose covariance matrix is $R_{\tilde{y}} = T^H R_y T$.

- Minimization of the output power

$$\min \tilde{w}^H R_{\tilde{y}} \tilde{w} \text{ subject to } \tilde{w}^H \tilde{a}_0 = 1 \quad \text{(PA-DF)}$$

  where $\tilde{a}_0 = T^H a_0$.

- The solution is given by

$$\tilde{w} = \alpha R_{\tilde{y}}^{-1} \tilde{a}_0 \Rightarrow w_{\text{PA-DF}} = \alpha T \left( T^H R_y T \right)^{-1} T^H a_0$$
Expression of the partially adaptive beamformer

GSC form

- New snapshots after transformation $\tilde{z}(k) = U^H z(k) = U^H B^H y(k)$ whose covariance matrix is $R_{\tilde{z}} = U^H R_z U$.

- Minimization of the output power

$$\min_{\tilde{w}_a} \mathcal{E} \left\{ |d(k) - \tilde{w}_a^H \tilde{z}(k)|^2 \right\} \quad \text{(PA-GSC)}$$

- The solution is given by

$$\tilde{w}_a = R_{\tilde{z}}^{-1} r_{d\tilde{z}} = \left( U^H R_z U \right)^{-1} U^H r_{dz}$$

$$w_{PA-GSC} = w_{CBF} - BUR_{\tilde{z}}^{-1} r_{d\tilde{z}}$$
Analysis of the partially adaptive MVDR

**SINR loss for fixed $T$ and $R_y = C$**

The SINR loss of the partially adaptive beamformer $w = T\tilde{w} = \alpha T\tilde{C}^{-1}\tilde{a}_0$ with fixed $T$ is distributed according to

$$\rho_{PA-MVDR} = \frac{d}{a} \left[ 1 + \frac{\chi^2_{2R}(0)}{\chi^2_{2(K-R+1)}(0)} \right]^{-1}$$

where

$$a = \frac{\tilde{a}_0^H \tilde{C}^{-1} \tilde{a}_0}{a_0^H C^{-1} a_0} = \frac{a_0^H T (T^H C T)^{-1} T^H a_0}{a_0^H C^{-1} a_0}$$

$$= \frac{\text{energy of } C^{-1/2}a_0 \text{ in } R \{ C^{1/2} T \}}{\text{energy of } C^{-1/2}a_0}$$
Analysis of the partially adaptive MVDR

⇒ partially adaptive beamforming is potentially very effective in low sample support, provided that \( T \) is well chosen.
Selection of matrices $T$ and $U$

**Fixed transformations**

- For instance using subarrays or spatial filtering, i.e.

\[
T = \begin{bmatrix} a(\tilde{\theta}_1) & a(\tilde{\theta}_2) & \cdots & a(\tilde{\theta}_R) \end{bmatrix} \\
U = B^H \begin{bmatrix} a(\tilde{\theta}_1) & a(\tilde{\theta}_2) & \cdots & a(\tilde{\theta}_R) \end{bmatrix}
\]

- In this case, the columns of $U$ can be viewed as beamformers aimed at intercepting the interference.

- Require some prior knowledge about the interference DOA in order for them to pass through the beams.
Selection of matrices $T$ and $U$

Value of $a$ when $U = B^H \begin{bmatrix} a(\tilde{\theta}_1) & a(\tilde{\theta}_2) \end{bmatrix}$ and $\tilde{\theta}_i$ drawn randomly around $\theta_i$:

![Graph 1](#)

$\tilde{\theta}_i \sim \mathcal{U}(\left[\theta_i - \frac{\theta_{sdb}}{2}, \theta_i + \frac{\theta_{sdb}}{2}\right])$

![Graph 2](#)

$\tilde{\theta}_i \sim \mathcal{U}(\left[\theta_i - \theta_{sdb}, \theta_i + \theta_{sdb}\right])$
Selection of matrices $T$ and $U$

Random transformations

- The idea\(^a\) is to use $L$ matrices $U_\ell$ drawn from a uniform distribution on the manifold of semi-unitary $(N - 1) \times R$ matrices, i.e.

$$U_\ell = X_\ell \left( X_\ell^H X_\ell \right)^{-H/2}; \quad X_\ell \overset{d}{=} \mathcal{CN} \left( 0, I_{N-1}, I_R \right)$$

and to average the corresponding weight vectors $\tilde{w}_\ell$, yielding

$$w = w_{CBF} - B \left[ \frac{1}{L} \sum_{\ell=1}^L U_\ell \left( U_\ell^H R_z U_\ell \right)^{-1} U_\ell^H r_{dz} \right]$$

$$= w_{CBF} - B \left[ \frac{1}{L} \sum_{\ell=1}^L X_\ell \left( X_\ell^H R_z X_\ell \right)^{-1} X_\ell^H r_{dz} \right]$$

Marzetta’s method based on random $\mathbf{U}$

The matrices $\mathbf{U}_\ell$ are drawn from a uniform distribution on the manifold of semi-unitary matrices, or from a Gaussian distribution $\mathcal{C}\mathcal{N}(0, \mathbf{I}_{N-1}, \mathbf{I}_R)$. 
Selection of matrices $T$ and $U$

**Adaptive transformations**

Matrices $T$ or $U$ depend on the snapshots. For example, in GSC form, if

$$R_z = \sum_{n=1}^{N-1} \lambda_n u_n u_n^H; \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N-1}$$

one can choose

- the $R$ principal eigenvectors of $R_z$ (Principal Component), i.e.

  $$U = [u_1 \cdots u_R] \Rightarrow w_{pc-gsc} = w_{CBF} - BU \Lambda^{-1} U^H r_{dz}$$

  where $\Lambda = \text{diag}\{\lambda_1, \cdots, \lambda_R\}$.

- the $R$ eigenvectors which contribute most to increasing the SINR (Cross Spectral Metric).
Partially adaptive beamforming: SINR versus $K$
Partially adaptive beamforming: SINR versus $K$
Partially adaptive beamforming: SINR versus $R$

![SINR of PC and CSM beamformers](image)

- Plot title: SINR of PC and CSM beamformers
- X-axis: Rank of the transformation
- Y-axis: dB
- Legend:
  - optimum
  - MVDR
  - PC
  - CSM

K=20

Beamforming
Beamforming: synthesis

- **Conventional beamforming** $\mathbf{w}_{CBF} = (\mathbf{a}_0^H \mathbf{a}_0)^{-1} \mathbf{a}_0$. *Optimal in white noise*, $\theta_{3dB} = 0.9 \left( \frac{N d}{\lambda} \right)^{-1}$, sidelobes at $-13$dB.

- **Adaptive beamforming** $\mathbf{w}_{opt} \propto \mathbf{C}^{-1} \mathbf{a}_s$, $\mathbf{w}_{MVDR} \propto \mathbf{C}^{-1} \mathbf{a}_0$, $\mathbf{w}_{MPDR} \propto \mathbf{R}^{-1} \mathbf{a}_0$
  
  - all equivalent if $\mathbf{R}$, $\mathbf{C}$ known and $\mathbf{a}_s = \mathbf{a}_0$
  - $\text{SINR}_{opt} \gtrsim \text{SINR}_{MVDR} \gg \text{SINR}_{MPDR}$ when $\mathbf{a}_s \neq \mathbf{a}_0$
  - $\text{SINR}_{MVDR-SMI} \gg \text{SINR}_{MPDR-SMI}$: convergence for about $2N$ snapshots for MVDR, $N \times \text{SINR}_{opt}$ for MPDR

- **Diagonal loading**: *helps to mitigate both finite-sample errors and steering vector errors*. Especially useful in MPDR context with low power signal of interest.

- **Partially adaptive beamforming**: enables one to achieve *faster convergence* by operating in low-dimensional subspace. Especially effective with strong, low-rank interference.
Contents

1 Introduction

2 Array processing model

3 Beamforming

4 Direction of arrival estimation
   Problem formulation
   Non parametric methods (beamforming)
   Parametric methods for DOA estimation
   Maximum likelihood estimation
   Subspace-based methods
   Covariance fitting
   Synthesis
The direction of arrival estimation problem

**Problem formulation**

Given a collection of $K$ snapshots which can possibly be modeled as

$$y(k) = \sum_{p=1}^{P} a(\theta_p) s_p(k) + n(k),$$

estimate the directions of arrival (DoA) $\theta_1, \ldots, \theta_P$:

![Diagram](image.png)

**Approaches**

- Non parametric approaches which do not necessarily rely on a model for $y(k)$: similar to Fourier-based methods in time domain;

- Parametric approaches where a model is assumed and its properties (algebraic structure, distribution) are exploited.
Beamforming for direction finding purposes

- The idea is to form a beam $w(\theta)$ for each angle $\theta$ and to evaluate the power $\mathcal{E} \left\{ \left| y_F(k) \right|^2 \right\} = \mathcal{E} \left\{ \left| w^H(\theta)y(k) \right|^2 \right\} = w^H(\theta)Rw(\theta)$ at the output of the beamformer versus $\theta$:

\[
\begin{array}{c}
y(k) \\
\downarrow \\
w(\theta)
\end{array} \rightarrow \quad P_w(\theta) = \mathcal{E} \left\{ \left| w^H(\theta)y(k) \right|^2 \right\}
\]

- Large peaks should provide the directions of arrival:
Conventional beamforming for direction finding purposes

**Conventional beamformer**

The conventional beamformer $\mathbf{w}_{CBF}(\theta) = \mathbf{a}(\theta)/N$ can be used, which yields the output power

$$
\mathbf{w}_{CBF}^H(\theta) \mathbf{R} \mathbf{w}_{CBF}(\theta) = N^{-2} \mathbf{a}^H(\theta) \mathbf{R} \mathbf{a}(\theta)
$$

**In practice**

With $K$ snapshots available, $\mathbf{R}$ is estimated as

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

and subsequently the output power as

$$
P_{CBF}(\theta) = N^{-2} \mathbf{a}^H(\theta) \hat{\mathbf{R}} \mathbf{a}(\theta)
$$

Direction of arrival estimation  Non parametric methods (beamforming)
CBF and Fourier analysis

• The estimated power at the output of the CBF writes

\[ P_{CBF}(\theta) = \frac{1}{N^2} a^H(\theta) \hat{R} a(\theta) \]

\[ = \frac{1}{KN^2} \sum_{k=1}^{K} |a^H(\theta)y(k)|^2 \]

\[ = \frac{1}{KN^2} \sum_{k=1}^{K} \left| \sum_{n=1}^{N} y_n(k)e^{-i2\pi(n-1)f} \right|^2 \]

where \( f = \frac{d}{\lambda} \sin \theta \).

• The inner sum is recognized as the (spatial) **Fourier transform** of each snapshot.
MPDR beamforming for direction finding purposes

Capon’s method

If the MPDR beamformer

$$w_{MPDR}(\theta) = \frac{R^{-1}a(\theta)}{a^H(\theta)R^{-1}a(\theta)}$$

is used, the output power then writes

$$w_{MPDR}^H(\theta)Rw_{MPDR}(\theta) = \frac{a^H(\theta)R^{-1}RR^{-1}a(\theta)}{[a^H(\theta)R^{-1}a(\theta)]^2} = \frac{1}{a^H(\theta)R^{-1}a(\theta)}$$

which in practice yields

$$P_{Capon}(\theta) = \frac{1}{a^H(\theta)\hat{R}^{-1}a(\theta)}$$
Comparison CBF-Capon (low resolution scenario)

\[
\theta = [-30^\circ, 10^\circ, 20^\circ]
\]

\[
\theta_{3dB} \sim 5.1^\circ
\]
Comparison CBF-Capon (high resolution scenario)

\[ \theta = [-30^\circ, 10^\circ, 15^\circ] \]

\[ \theta_{3dB} \sim 5.1^\circ \]

Direction of arrival estimation  Non parametric methods (beamforming)
Model-based methods

**Principle**

Based on the model

\[ y(k) = A(\theta)s(k) + n(k) \]

where \( \theta = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_P]^T \),

\[ A(\theta) = \begin{bmatrix} a(\theta_1) & a(\theta_2) & \cdots & a(\theta_P) \end{bmatrix} \]

\[ s(k) = \begin{bmatrix} s_1(k) & s_2(k) & \cdots & s_P(k) \end{bmatrix}^T \]

and \( a(\theta) \) stands for the steering vector.
Classes of methods

- **Maximum Likelihood methods** are based on maximizing the likelihood function, which amounts to finding the unknown parameters which make the observed data the more likely.

- **Subspace-based methods** rely on the fact that the signal subspace coincides with the subspace spanned by the principal eigenvectors of $\mathbf{R}$. Moreover, the latter is orthogonal to the subspace spanned by the minor eigenvectors. These two algebraic properties are exploited for direction finding.

- **Covariance matching** relies on a model $\mathbf{R}(\eta)$ for the covariance matrix and looks for the model parameters which minimize the distance between $\mathbf{R}(\eta)$ and the sample covariance matrix $\hat{\mathbf{R}}$. 
Maximum Likelihood Estimation

- The MLE consists in finding the parameter vector $\eta$ which maximizes the likelihood function $p(Y; \eta)$ of the snapshots $Y = [y(1), y(2), \cdots, y(k)]$, where $\eta$ is the model parameter vector.

😊 Asymptotically efficient.

😢 Multi-dimensional optimization problem (usually) $\Rightarrow$ computational complexity, possible convergence to local maxima.
Stochastic (unconditional) MLE

- Assume that $s(k)$ is Gaussian distributed with $\mathcal{E}\{s(k)\} = 0$, and a covariance matrix $R_s = \mathcal{E}\{s(k)s^H(k)\}$ which is full rank.
- The distribution of the snapshots is thus given by

  $$y(k) \sim \mathcal{CN}(0, R = A(\theta)R_sA^H(\theta) + \sigma^2 I)$$

- The likelihood function can be written as

  $$p(Y; \eta) = \prod_{k=1}^{K} \pi^{-N} |R|^{-1} e^{-y(k)^H R^{-1} y(k)}$$
Stochastic (unconditional) MLE

• The ML estimate is obtained as

\[
\hat{\eta} = \arg \min_{\theta, R_s, \sigma^2} - \log p(Y; \eta)
= \arg \min_{\theta, R_s, \sigma^2} \log |R| + \text{Tr} \left\{ R^{-1} \hat{R} \right\}
\]

• Closed-form solutions for \( \sigma^2 \) and \( R_s \) can be obtained so that the likelihood function is concentrated, yielding a minimization over the angles only:

\[
\hat{\theta}^{\text{sto}} = \arg \min_{\theta} \log \left| A(\theta) \hat{R}_s(\theta) A^H(\theta) + \hat{\sigma}^2(\theta) I \right|
\]
Deterministic (conditional) MLE

- The signal waveforms are assumed deterministic so that

\[ y(k) \sim CN(A(\theta)s(k), \sigma^2 I) \]

- The MLE is now given by

\[
\hat{\eta} = \arg \min_{\theta, s(k), \sigma^2} NK \log \sigma^2 + \sigma^{-2} \sum_{k=1}^{K} \| y(k) - A(\theta)s(k) \|^2
\]

- The likelihood function can be concentrated with respect to all \( s(k) \) and \( \sigma^2 \), and finally

\[
\hat{\theta}^{\text{det}} = \arg \min_{\theta} \text{Tr} \left\{ P_A(\theta) \hat{R} \right\}
\]

- For a single source \( \hat{\theta}^{\text{det}} = \arg \max_{\theta} \frac{1}{N} a^H(\theta) \hat{R} a(\theta) \equiv \text{CBF} \).
Subspace-based methods

Eigenvalue decomposition of the covariance matrix

If $P$ signals are present, one has

$$
\mathcal{R}\{\mathbf{A}(\theta)\}, \text{rank}=P
$$

$$
\mathbf{R} = \mathbf{A}(\theta)\mathbf{R}_s\mathbf{A}^H(\theta) + \sigma^2 \mathbf{I} \quad (\mathbf{R}_s \text{ assumed full-rank})
$$

$$
= \sum_{p=1}^{P} \lambda_p \mathbf{u}_p \mathbf{u}_p^H + 0 \sum_{p=P+1}^{N} \mathbf{u}_p \mathbf{u}_p^H + \sigma^2 \mathbf{I}
$$

$$
= \sum_{p=1}^{P} \lambda_p \mathbf{u}_p \mathbf{u}_p^H + 0 \sum_{p=P+1}^{N} \mathbf{u}_p \mathbf{u}_p^H + \sigma^2 \sum_{p=1}^{N} \mathbf{u}_p \mathbf{u}_p^H
$$

$$
= \mathbf{U}_s \boldsymbol{\Lambda}_s \mathbf{U}_s^H + \sigma^2 \mathbf{U}_n \mathbf{U}_n^H
$$

where $\mathbf{U}_s = [\mathbf{u}_1 \ldots \mathbf{u}_P] \perp \mathbf{U}_n = [\mathbf{u}_{P+1} \ldots \mathbf{u}_N]$. 
Subspace-based methods

Signal and noise subspaces

- \( \mathcal{R} \{ U_s \} = \mathcal{R} \{ A(\theta) \} : \) the signal subspace is spanned by \( U_s \) and hence \( U_s = A(\theta)T \) for some non-singular matrix \( T \).
- \( \mathcal{R} \{ U_n \} \) is orthogonal to \( \mathcal{R} \{ U_s \} = \mathcal{R} \{ A(\theta) \} \Rightarrow A^H(\theta)U_n = 0 \).

⇒ Subspace-based methods rely on either \( \bigcirc \) or \( \bigcirc \).
**MUSIC**

- **The signal steering vectors are orthogonal to** $U_n$

\[ U_n^H a(\theta_p) = 0 \iff u_n^H a(\theta_p) \text{ for } n = P + 1, \ldots, N \]

- One looks for the $P$ largest maxima of

\[ P_{\text{MUSIC}}(\theta) = \frac{1}{a^H(\theta) \hat{U}_n \hat{U}_n^H a(\theta)} = \frac{1}{\sum_{n=P+1}^{N} |a^H(\theta)u_n|^2} \]

on the rationale that, as $K$ grows large, $\hat{U}_n \to U_n$ and hence $P_{\text{MUSIC}}(\theta_p) \to \infty$.

- Many variants around MUSIC, e.g., SSMUSIC (Mc Cloud & Scharf).
Root-MUSIC

- Let \( a(z) = [1 \ z \ \cdots \ z^{N-1}]^T \). For a ULA, one can compute the \( P \) roots of
  \[
P_{\text{MUSIC}}(z) = a^T(z^{-1}) \hat{U}_n \hat{U}_n^H a(z)
  \]
  closest to the unit circle. The reason is that
  \[
a^T(e^{i2\pi \frac{d}{\lambda} \sin \theta_p}) \hat{U}_n \hat{U}_n^H a(e^{i2\pi \frac{d}{\lambda} \sin \theta_p}) = a^H(\theta_p) \hat{U}_n \hat{U}_n^H a(\theta_p) = 0
  \]

- \( P_{\text{MUSIC}}(z) = \sum_{n=-(N-1)}^{N-1} p_n z^{-n} \) has \( 2(N-1) \) roots, \( (N-1) \) of which inside the unit circle since
  \[
P_{\text{MUSIC}}(1/z^*) = a^T(z^*) \hat{U}_n \hat{U}_n^H a(1/z^*)
  = a^H(z) \hat{U}_n \hat{U}_n^H a^*(z^{-1})
  = a^T(z^{-1}) \hat{U}_n \hat{U}_n^H a(z)
  = P_{\text{MUSIC}}(z)
  \]
Low-resolution scenario

Eigenvalues of the covariance matrix

Direction of arrival estimation
Subspace-based methods
Low-resolution scenario

\[ \theta = [-30^\circ, 10^\circ, 20^\circ] \]

\[ \theta_{MB} \sim 5.1^\circ \]
High-resolution scenario

Eigenvalues of the covariance matrix

Direction of arrival estimation
Subspace-based methods
High-resolution scenario

Comparison CBF–MUSIC

$\theta = [-30^\circ, 10^\circ, 13^\circ]$

$\theta_{3\text{dB}} \sim 5.1^\circ$

Direction of arrival estimation
Subspace-based methods
Min-norm

- Let \( \mathbf{d} = \mathbf{U}_n \eta = [d_0 \ d_1 \ \cdots \ d_{N-1}]^T \) be an arbitrary vector in the noise subspace.

- Since \( \mathbf{d} \perp \mathbf{a}(\theta_p) \), \( D(z) = \sum_{n=0}^{N-1} d_n z^{-n} \) has \( P \) of its roots equal to \( e^{i2\pi \frac{d}{X} \sin \theta_p} \) and hence can serve to estimate \( \theta_p \).

- The min-norm method searches the vector in \( \mathcal{R} \{ \mathbf{U}_n \} \) with minimal norm. To avoid \( \mathbf{d} = 0 \), one considers

\[
\min_{\mathbf{d} \in \mathcal{R} \{ \mathbf{U}_n \}} \| \mathbf{d} \|^2 \quad \text{s. t.} \quad d_0 = 1 \Leftrightarrow \min_{\eta} \| \eta \|^2 \quad \text{s. t.} \quad \eta^H \mathbf{U}_n^H \mathbf{e}_1 = 1
\]

where \( \mathbf{e}_1 = [1 \ 0 \ 0 \ \cdots \ 0]^T \). The solution is

\[
\eta^*_n = \frac{\mathbf{U}_n^H \mathbf{e}_1}{\mathbf{e}_1^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{e}_1} \Rightarrow \mathbf{d}_{\text{Min-Norm}} = \frac{\mathbf{U}_n \mathbf{U}_n^H \mathbf{e}_1}{\mathbf{e}_1^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{e}_1}
\]
• Let us partition $A = A(\theta)$ as

$$A = \begin{bmatrix} A_1 \\ - \end{bmatrix} = \begin{bmatrix} - \\ A_2 \end{bmatrix}$$

where $A_1$ [resp. $A_2$] contains all but the last [resp. first] row of $A$.

• Then, for a ULA, we have

$$A_2 = A_1 \Phi; \quad \Phi = \text{diag}\left(\left\{ e^{i2\pi \frac{d}{\lambda} \sin \theta_p} \right\}_{p=1}^{P}\right) \quad (1)$$

• $\Phi$ conveys the useful information and can be deduced from $(A_1, A_2)$. The latter are unknown but $U_s = AT \Rightarrow$ can we find a similar relation for $U_s$?
ESPRIT

• Let us partition $\mathbf{U}_s$ as $\mathbf{A}$, i.e. $\mathbf{U}_s = \begin{bmatrix} \mathbf{U}_{s1} \\ - \end{bmatrix} = \begin{bmatrix} - \\ \mathbf{U}_{s2} \end{bmatrix}$. Then

$$\mathbf{U}_s = \mathbf{A}\mathbf{T} \Rightarrow \begin{cases} \mathbf{U}_{s1} = \mathbf{A}_1\mathbf{T} \\ \mathbf{U}_{s2} = \mathbf{A}_2\mathbf{T} = \mathbf{A}_1\mathbf{\Phi}\mathbf{T} \end{cases}$$

$$\Rightarrow \mathbf{U}_{s2} = \mathbf{U}_{s1}\mathbf{T}^{-1}\mathbf{\Phi}\mathbf{T}$$

$$\Rightarrow \mathbf{U}_{s2} = \mathbf{U}_{s1}\mathbf{\Psi}$$

• $\mathbf{\Psi}$ and $\mathbf{\Phi}$ share the same eigenvalues since $\mathbf{\Phi}\mathbf{u} = \lambda\mathbf{u}$ implies that $\mathbf{\Psi}\mathbf{T}^{-1}\mathbf{u} = \mathbf{T}^{-1}\mathbf{\Phi}\mathbf{u} = \lambda\mathbf{T}^{-1}\mathbf{u}$.

• It follows that the eigenvalues of $\mathbf{\Psi}$ are $\left\{ e^{i2\pi \frac{d}{\lambda} \sin \theta_p} \right\}_{p=1}^{P}$.

• In practice one solves in a least-squares sense $\hat{\mathbf{U}}_{s2} = \hat{\mathbf{U}}_{s1}\mathbf{\Psi}$ and computes the eigenvalues of $\hat{\mathbf{\Psi}}$. 

Direction of arrival estimation  Subspace-based methods
Subspace Fitting

- Since $\mathcal{R}\{\mathbf{U}_s\} = \mathcal{R}\{\mathbf{A}(\theta)\}$, there exists a full-rank matrix $\mathbf{T} (P \times P)$ such that

\[
\mathbf{U}_s = \mathbf{A}(\theta)\mathbf{T}
\]

- The idea is to look for the DOA which minimize the error between the subspaces spanned by $\hat{\mathbf{U}}_s$ and $\mathbf{A}(\theta)$:

\[
\hat{\theta}, \hat{T} = \arg\min_{\theta, T} \left\| \hat{\mathbf{U}}_s - \mathbf{A}(\theta)\mathbf{T} \right\|^2_W
\]

\[
= \arg\min_{\theta, T} \text{Tr} \left\{ \left[ \hat{\mathbf{U}}_s - \mathbf{A}(\theta)\mathbf{T} \right] \mathbf{W} \left[ \hat{\mathbf{U}}_s - \mathbf{A}(\theta)\mathbf{T} \right]^H \right\}
\]
Subspace Fitting

- There exists a closed-form solution for $T$ and finally

$$
\hat{\theta}^{\text{SSF}} = \arg \min_{\theta} \text{Tr} \left\{ P_A(\theta) \hat{U}_s W \hat{U}_s^H \right\}
$$

- Alternative: use the fact that

$$
\mathcal{R} \{ \mathbf{U}_n \} = \mathcal{N} \left\{ \mathbf{A}^H(\theta) \right\} \Rightarrow \mathbf{U}_n^H \mathbf{A}(\theta) = \mathbf{0}
$$

and estimate the angles as

$$
\hat{\theta}^{\text{NSF}} = \arg \min_{\theta} \left\| \mathbf{U}_n^H \mathbf{A}(\theta) \right\|_W^2
$$
Covariance fitting

- The covariance matrix is given by $R(\theta, P, \sigma) = R_s(\theta, P) + Q(\sigma)$

$$r = \text{vec}(R) = \Psi(\theta)P + \Sigma\sigma = \begin{bmatrix} \Psi(\theta) & \Sigma \end{bmatrix} \begin{bmatrix} P \\ \sigma \end{bmatrix} \triangleq \Phi(\theta)\alpha$$

- The parameters are estimated by minimizing the error between $R$ and its estimate $\hat{R}$:

$$\hat{\theta}, \hat{\alpha} = \arg \min [\hat{r} - \Phi(\theta)\alpha] W^{-1} [\hat{r} - \Phi(\theta)\alpha]$$

- The criterion can be concentrated with respect to $\alpha$: minimization with respect to $\theta$ only.
Covariance fitting

- In case of independent Gaussian distributed snapshots, \( W_{\text{opt}} = R^T \otimes R \) and covariance matching estimates are asymptotically (i.e. when \( K \to \infty \)) equivalent to ML estimates.
- In contrast to MLE, no need for assumptions on the pdf, only an assumption on \( R \). The criterion is usually simpler to minimize.
- Covariance matching can be used with full-rank covariance matrix \( R_s \) while subspace methods require the latter to be rank deficient.
## Synthesis

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Algorithm</th>
<th>Performance</th>
<th>Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML</td>
<td>distribution</td>
<td>optimization</td>
<td>optimal</td>
</tr>
<tr>
<td>COMET</td>
<td>$\mathbb{R}$</td>
<td>optimization</td>
<td>$\approx$ optimal</td>
</tr>
<tr>
<td>MUSIC</td>
<td>$\mathbb{R}$</td>
<td>EVD</td>
<td>$\approx$ optimal</td>
</tr>
</tbody>
</table>

**Direction of arrival estimation**

**Synthesis**
Conclusions

- Array processing, thanks to additional degrees of freedom, enables one to perform spatial filtering of signals.
- Adaptive beamforming, possibly with reduced-rank transformations, enables one to achieve high SINR with a fast rate of convergence in adverse conditions (interference, noise).
- Robustness issues are of utmost importance in practical systems, and should be given a careful attention.
- Non-parametric direction finding methods are simple and robust but may suffer from a lack of resolution.
- Parametric methods offer high resolution, often at the price of degraded robustness.
References