Using random matrix theory in array processing applications: the small sample size regime

Xavier Mestre

xavier.mestre@cttc.cat

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)

Parc Mediterrani de la Tecnologia, Castelldefels, Barcelona, Spain

The third EPFL-UMLV Workshop on Random Matrices, Information Theory and Applications, 3-4 December 2009

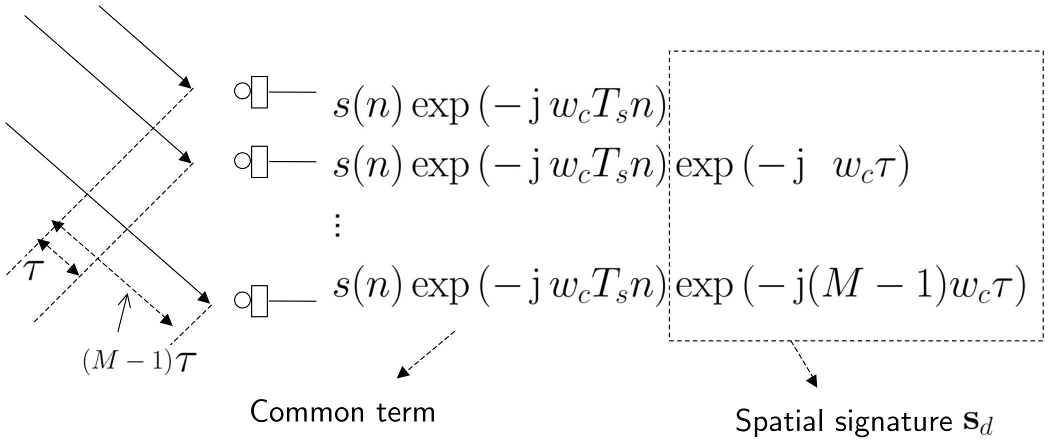
Outline

- Introduction to array signal processing.
- Direction of arrival estimation and signal power determination. The ML approach.
- The MUSIC approach: asymptotic behavior of the eigenvectors of the sample covariance matrix.
- Estimating with random matrix theory. N-consistency versus M, N-consistency. Estimation of the eigenvectors of the sample covariance matrix.
- Estimating the power of the sources and the eigenvalues of the sample covariance matrix.

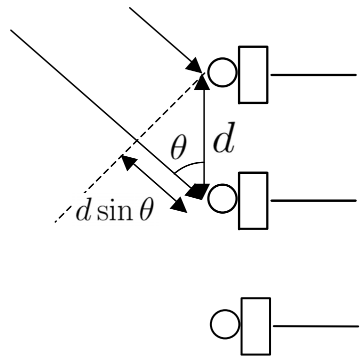
Array signal processing: introduction

Assume that we receive K narrowband signals from a receiver equipped with an antenna array of M elements:

$$\mathbf{y}(n) = \sum_{k=1}^{K} s_k(n)\mathbf{a}(\theta_k) + \mathbf{n}(n) \Longrightarrow \boxed{\mathbf{y}(n) = \mathbf{A}(\Theta)\,\mathbf{s}(n) + \mathbf{n}(n)} \qquad \mathbf{A}(\Theta) = \begin{bmatrix} \mathbf{a}(\theta_1) & \mathbf{a}(\theta_2) & \cdots & \mathbf{a}(\theta_K) \end{bmatrix}$$



$$\boxed{\tau = \frac{d\sin\theta}{\lambda_c w_c}}$$



Signal models in array processing

Traditionally, two different assumptions are used in array signal processing, giving rise to two different signal models.

$$\mathbf{y}(n) = \mathbf{A}(\Theta)\,\mathbf{s}(n) + \mathbf{n}(n)$$

- The **conditional or deterministic signal model**: the collection of signal vectors $\mathbf{s}(n)$ is treated as an unknown deterministic parameter. This model is sometimes referred to as the information-plus-noise model.
- The **unconditional or random signal model**: the collection of signal vectors $\mathbf{s}(n)$ is treated as a random parameter, usually having zero mean, circular symmetry and covariance

$$\Phi_S = \mathbb{E}\left[\mathbf{s}(n)\mathbf{s}^H(n)\right], \quad \mathbb{E}\left[\mathbf{s}(n)\mathbf{s}^T(n)\right] = \mathbf{0}.$$

This presentation will only focus on the random (or unconditional) signal model.

Problem statement

In this talk, we will consider the following problem. Consider a set of K sources impinging on an array of M sensors/antennas. We will assume that $1 \leq K < M$.

We work with a fixed number of snapshots N,

$$\{\mathbf{y}(1), \ldots, \mathbf{y}(N)\}$$

assumed i.i.d., with zero mean and covariance $\mathbf{R}_M = \mathbf{A}(\Theta)\,\Phi_S \mathbf{A}(\Theta)^H + \sigma^2 \mathbf{I}_M$. Assume that from these snapshots, we construct the sample covariance matrix

$$\hat{\mathbf{R}}_M = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n) = \mathbf{R}_M^{1/2}\left(\frac{1}{N}\sum_{n=1}^{N} \mathbf{u}(n)\mathbf{u}^H(n)\right)\mathbf{R}_M^{1/2}.$$

Problem statement: knowing the number of sources present in the scenario (K) and the array manifold $\mathbf{a}(\theta)$, determine the directions of arrival of these sources, i.e. $\Theta = [\theta_1, \ldots, \theta_k]$ and their power $\{\Phi_S\}_{kk}$ from the sample covariance matrix, $\hat{\mathbf{R}}_M$

The (Gaussian) Maximum Likelihood estimator

Let us now consider the maximum likelihood estimator of the above problem, departing from the assumption that

$$\mathbf{y}(n) = \sum_{k=1}^{K} s_k(n)\mathbf{a}\left(\theta_k\right) + \mathbf{n}(n) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_M\right).$$

The estimator is obtained by maximizing the log-likelihood fuction with respect to all the unknowns, namely

$$\hat{\Theta} = \arg_1 \min_{\Theta, \Phi_S, \sigma^2} \left\{ \log\det\left(\mathbf{R}_M\right) + \operatorname{tr}\left[\mathbf{R}_M^{-1}\hat{\mathbf{R}}_M\right] \right\}$$

where we recall that $\mathbf{R}_M = \mathbf{A}\left(\Theta\right)\Phi_S\mathbf{A}\left(\Theta\right)^H + \sigma^2\mathbf{I}_M$. We assume that $M < N$ and that $\mathbf{A}\left(\Theta\right)$ is full column rank whenever the DoAs are different.

The estimator can be further described by inserting the estimation of the nuisance parameters (Φ_S, σ^2) back into the cost function.

The (Gaussian) Maximum Likelihood estimator (II)

By minimizing with respect to the signal covariance matrix, we obtain

$$\hat{\Phi}_S = \left[\mathbf{A}^H(\Theta)\,\mathbf{A}(\Theta)\right]^{-1}\mathbf{A}(\Theta)^H\,\hat{\mathbf{R}}_M\mathbf{A}(\Theta)\left[\mathbf{A}^H(\Theta)\,\mathbf{A}(\Theta)\right]^{-1} - \sigma^2\left[\mathbf{A}^H(\Theta)\,\mathbf{A}(\Theta)\right]^{-1}$$

and replacing this into the ML cost function we obtain

$$\hat{\sigma}^2 = \frac{1}{M-K}\operatorname{tr}\left[\mathbf{P}_{\mathbf{A}}^{\perp}(\Theta)\,\hat{\mathbf{R}}_M\right]$$

Finally replacing the noise power estimation in the log-likelihood cost function,

$$\hat{\Theta} = \arg\min_{\Theta}\eta(\Theta) \qquad \eta(\Theta) = \frac{1}{M}\log\det\left(\mathbf{P}_A(\Theta)\,\hat{\mathbf{R}}_M\mathbf{P}_A(\Theta) + \hat{\sigma}^2\mathbf{P}_A^{\perp}(\Theta)\right)$$

Advantages: high accuracy, high resolution even in low SNR and low sample volumes.

Disadvantage: the multidimensionals search is computationally very expensive, assumes Gaussian statistics.

The MUSIC approach

The eigendecomposition of $\mathbf{R}_M$ allows us to differentiate between signal and noise subspaces:

$$\mathbf{R}_M = \left[\begin{array}{cc} \mathbf{E}_S & \mathbf{E}_N \end{array} \right] \left[\begin{array}{cc} \mathbf{\Lambda}_S & \mathbf{0} \\ \mathbf{0} & \sigma^2 \mathbf{I}_{M-K} \end{array} \right] \left[\begin{array}{cc} \mathbf{E}_S & \mathbf{E}_N \end{array} \right]^H .$$

It turns out that $\mathbf{E}_N^H \mathbf{a}(\theta_k) = \mathbf{0}$, $k = 1 \ldots K$.

Since $\mathbf{R}_M$ is unknown, one must work with the sample covariance matrix

$$\hat{\mathbf{R}}_M = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n) = \left[\begin{array}{cc} \hat{\mathbf{E}}_S & \hat{\mathbf{E}}_N \end{array} \right] \left[\begin{array}{cc} \hat{\mathbf{\Lambda}}_S & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Lambda}}_N \end{array} \right] \left[\begin{array}{cc} \hat{\mathbf{E}}_S & \hat{\mathbf{E}}_N \end{array} \right]^H$$

The MUSIC algorithm uses the sample noise eigenvectors, and searches for the deepest local minima of the cost function

$$\eta_{\mathsf{MUSIC}}(\theta) = \mathbf{a}^H(\theta)\hat{\mathbf{E}}_N \hat{\mathbf{E}}_N^H \mathbf{a}(\theta) .$$

It is interesting to investigate the behavior of $\eta_{\mathsf{MUSIC}}(\theta)$ when M, N have the same order of magnitude (low sample regime).

Asymptotic behavior of MUSIC

The MUSIC algorithm suffers from the **breakdown effect**. The performance breaks down when the number of samples or the SNR falls below a **certain threshold**. Cause: $\hat{\mathbf{E}}_N$ is not a very good estimator of $\mathbf{E}_N$ when M, N have the same order of magnitude.

The performance breakdown effect can be easily analyzed using random matrix theory, (under a noise **eigenvalue separation assumption**): $|\eta_{\text{MUSIC}}(\theta) - \bar{\eta}_{\text{MUSIC}}(\theta)| \to 0$

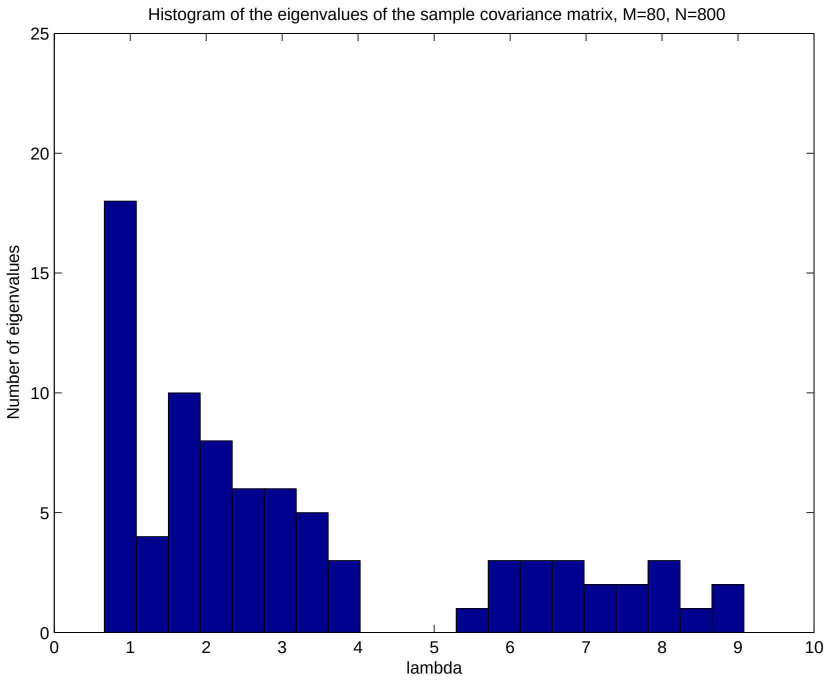
$$\bar{\eta}_{\text{MUSIC}}(\theta) = \mathbf{a}^H(\theta) \left(\sum_{k=1}^{M} w(k) \mathbf{e}_k \mathbf{e}_k^H \right) \mathbf{a}(\theta)$$

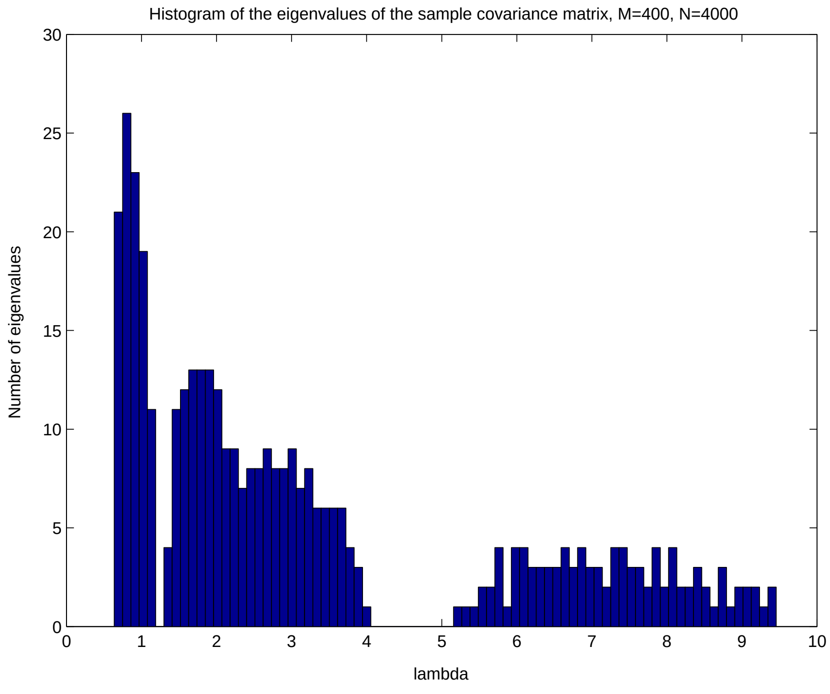
$$w(k) = \begin{cases} 1 - \frac{1}{M-K} \sum_{r=M-K+1}^{M} \left(\frac{\sigma^2}{\lambda_r - \sigma^2} - \frac{\mu_1}{\lambda_r - \mu_1} \right) & k \leq M-K \\ \frac{\sigma^2}{\lambda_k - \sigma^2} - \frac{\mu_1}{\lambda_k - \mu_1} & k > M-K \end{cases}$$

where $\{\mu_r,\ r = 1, \ldots, M\}$ are the solutions to $\frac{1}{M} \sum_{r=1}^{M} \frac{\lambda_r}{\lambda_r - \mu} = \frac{1}{c}$.

The sample covariance matrix

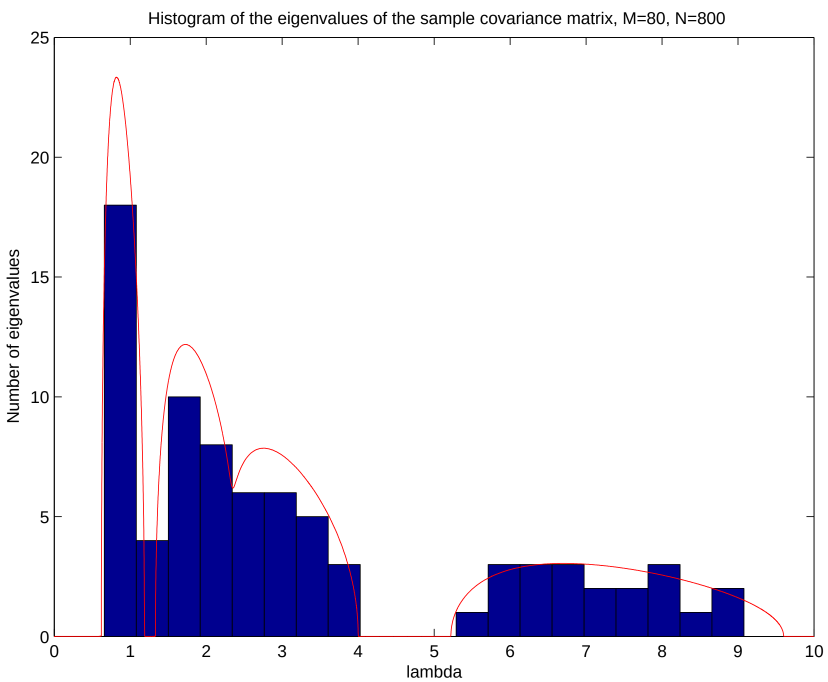
We assume that we collect N independent samples (snapshots) from an array of M antennas: $\hat{\mathbf{R}}_M = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}^H(n)$. Example: $\mathbf{R}_M$ has 4 eigenvalues $\{1, 2, 3, 7\}$ with equal multiplicity.

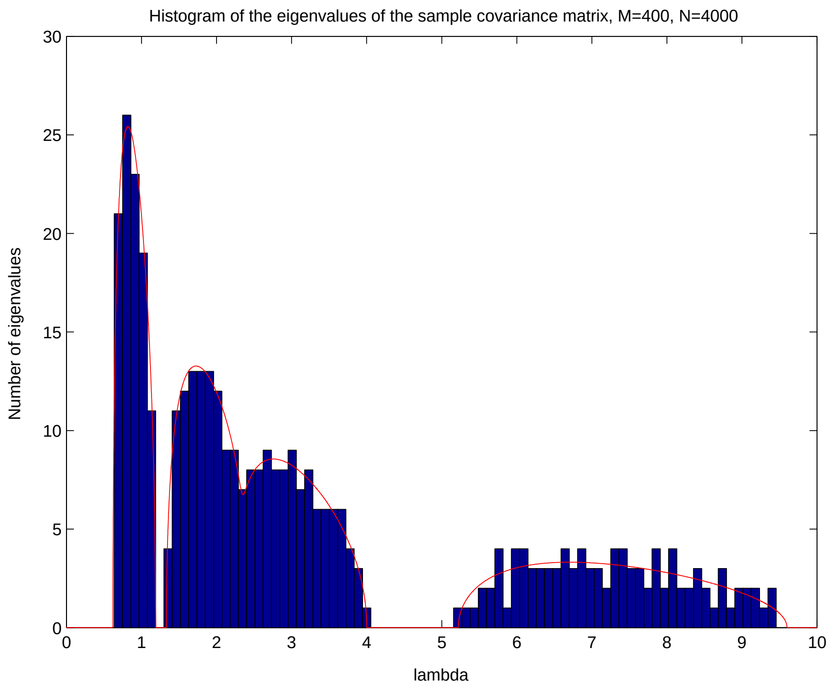




The sample covariance matrix

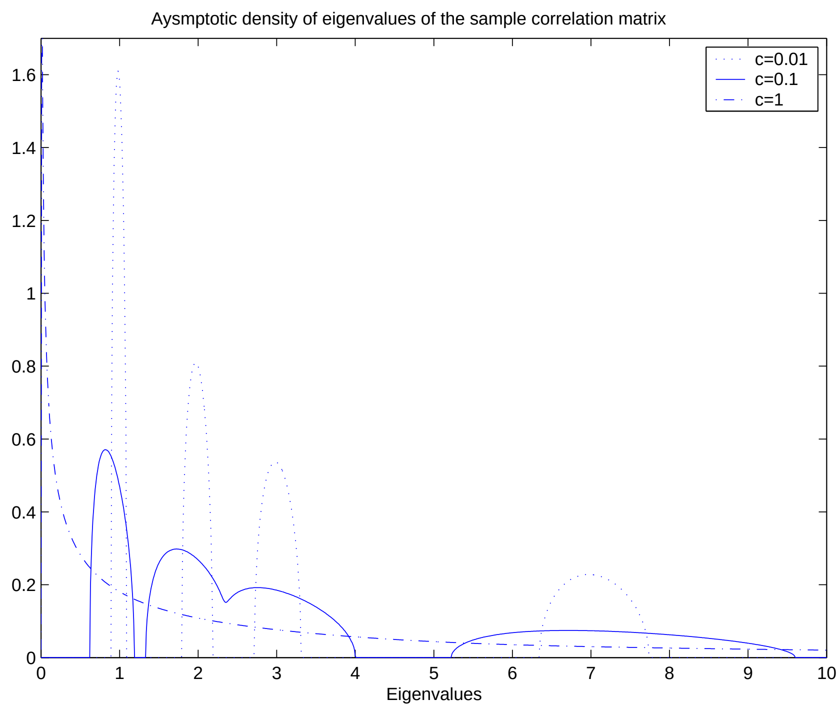
We assume that we collect N independent samples (snapshots) from an array of M antennas: $\hat{\mathbf{R}}_M = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}^H(n)$. Example: $\mathbf{R}_M$ has 4 eigenvalues $\{1,2,3,7\}$ with equal multiplicity.





The sample covariance matrix: asymptotic properties

When both $M, N \to \infty$, $M/N \to c$, $0 < c < \infty$, the e.d.f. of the eigenvalues of $\hat{\mathbf{R}}_M$ tends to a deterministic density function. Example: $\mathbf{R}_M$ has 4 eigenvalues $\{1, 2, 3, 7\}$ with equal multiplicity.



Convergence of the eigenvalues and eigenvectors of the sample covariance matrix

- Can we characterize the asymptotic sample eigenvalue behavior analytically? What about the eigenvectors?
- We will consider the following two quantities:

$$\hat{b}_M(z) = \frac{1}{M} \operatorname{tr}\left[\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\right], \quad \hat{m}_M(z) = \mathbf{a}^H \left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1} \mathbf{b}$$

  where $z \in \mathbb{C}^+$ and $\mathbf{a}, \mathbf{b} \in \mathbb{C}^M$.
- Note that $\hat{b}_M(z)$ depends only on the eigenvalues of $\hat{\mathbf{R}}_M$, whereas $\hat{m}_M(z)$ depends on both the eigenvalues and the eigenvectors of this matrix.
- Much of the asymptotic behavior of the eigenvalues and eigenvectors of $\hat{\mathbf{R}}_M$ can be inferred from this two quantities.

Convergence of the eigenvalues and eigenvectors of the sample covariance matrix

Assuming that $\mathbf{u}(n)$ have zero mean, circular symmetry, finite 8th moment,

$$\left|\hat{b}_M(z) - \bar{b}_M(z)\right| \to 0, \quad \left|\hat{m}_M(z) - \bar{m}_M(z)\right| \to 0$$

almost surely for all $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}\,[z] > 0\}$ as $M, N \to \infty$ at the same rate, where $\bar{b}_M(z) = b$ is the unique solution to the following equation in the set $\{b \in \mathbb{C} : -(1-c)/z + cb \in \mathbb{C}^+\}$:

$$b = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\lambda_m (1 - c - czb) - z}$$

and

$$\bar{m}_M(z) = \frac{f_M(z)}{z} \mathbf{a}^H \left(\mathbf{R}_M - f_M(z)\mathbf{I}_M\right)^{-1} \mathbf{b}, \quad f_M(z) = \frac{z}{1 - c - cz\bar{b}_M(z)}.$$

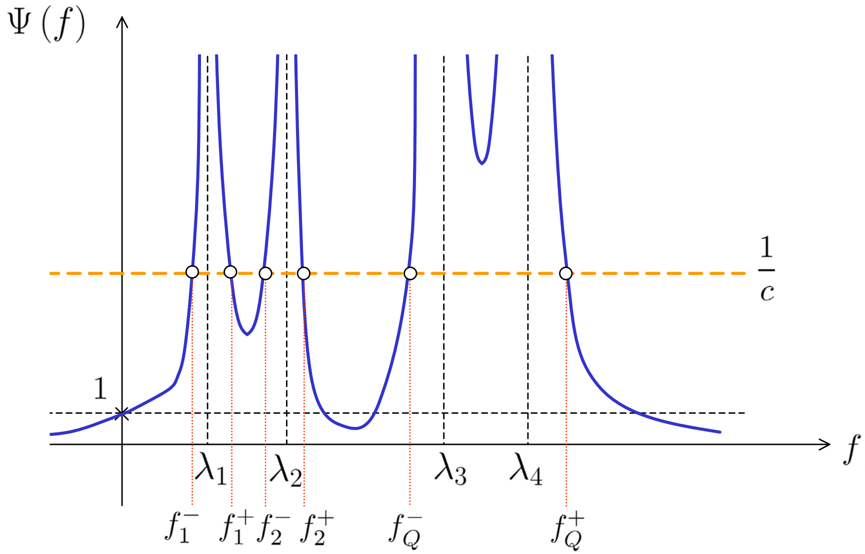
If $q_M(x)$ is the asymptotic sample eigenvalue *density*, then

$$q_M(x) = \lim_{y \to 0^+} \mathrm{Im}\left[\bar{b}_M(x + \mathrm{j}\, y)\right].$$

Determining the support of $q_M(x)$

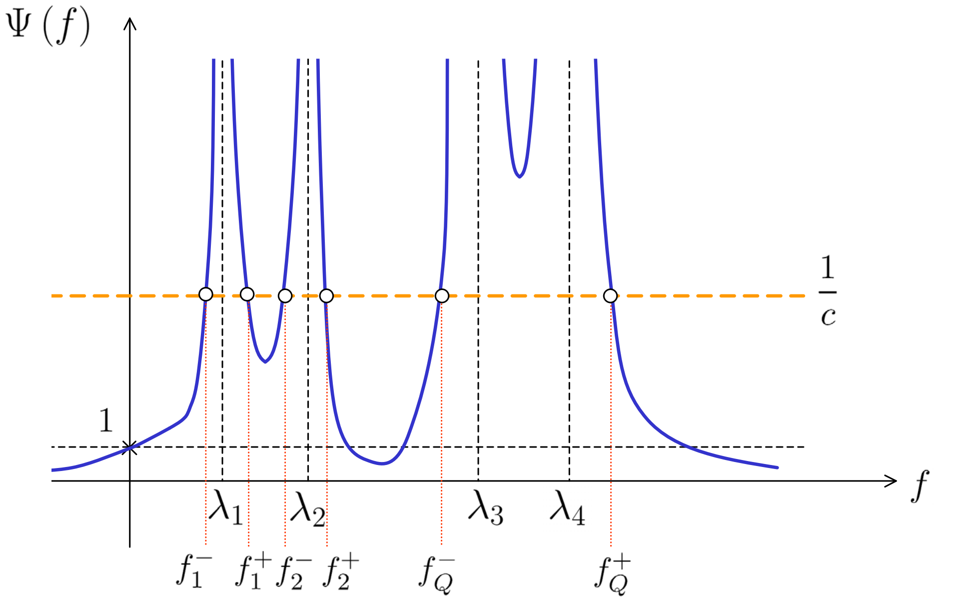
Consider the following equation in f, which has $2Q$ solutions (counting multiplicities)

$$\Psi\left(f\right) = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m - f}\right)^2 = \frac{1}{c}.$$



Determining the support of $q_M(x)$

The solutions are denoted $\left\{f_1^-, f_1^+, \ldots, f_Q^-, f_Q^+\right\}$. For each eigenvalue λ_m, there exists a single $q \in \{1 \ldots Q\}$, such that $\lambda_m \in \left(f_q^-, f_q^+\right)$. In other words, each eigenvalue is associated with a single cluster, but one cluster may be associated with multiple eigenvalues.



Determining the support of $q_M(x)$

It turns out that Q is the number of clusters in $q(x)$, and the position of each cluster is given by

$$x_q^+ = \Phi(f_q^+), \quad x_q^- = \Phi(f_q^-)$$

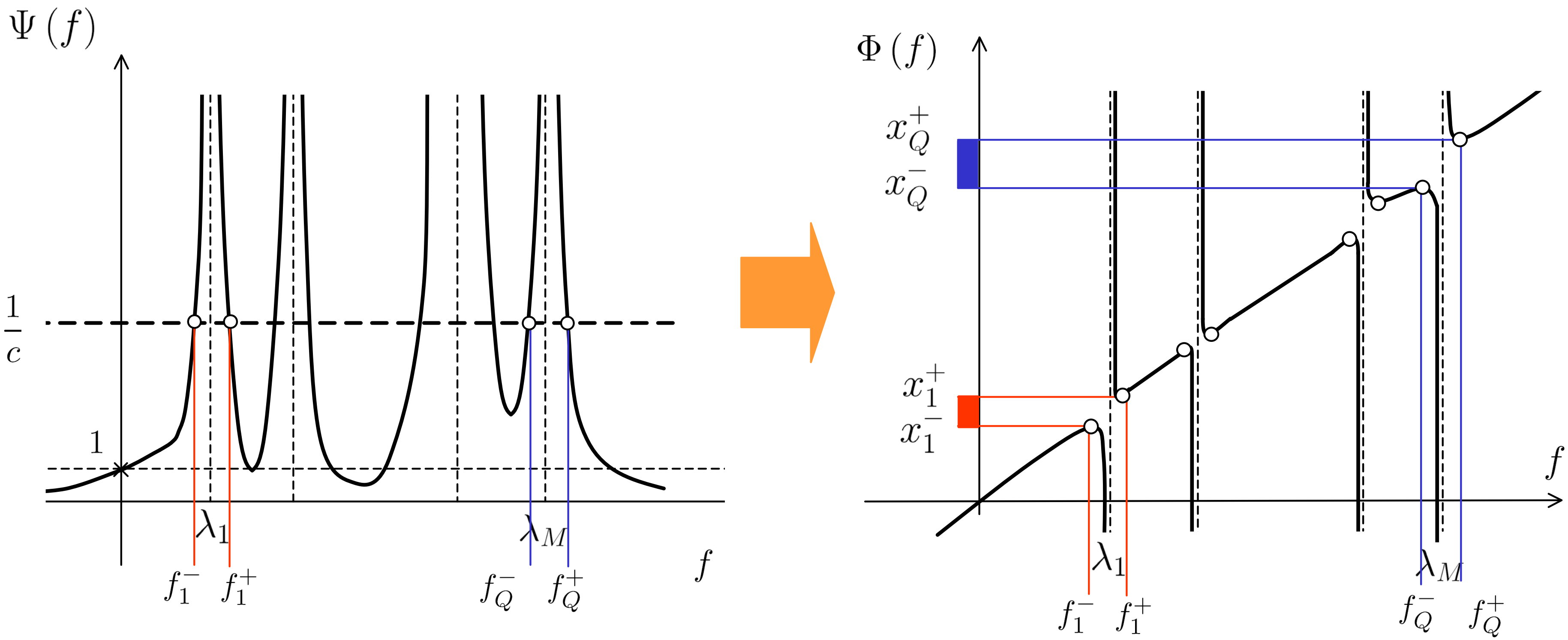
where the function $\Phi(f)$ is defined as

$$\Phi(f) = f\left(1 - \frac{c}{M}\sum_{m=1}^{M}\frac{\lambda_m}{\lambda_m - f}\right).$$

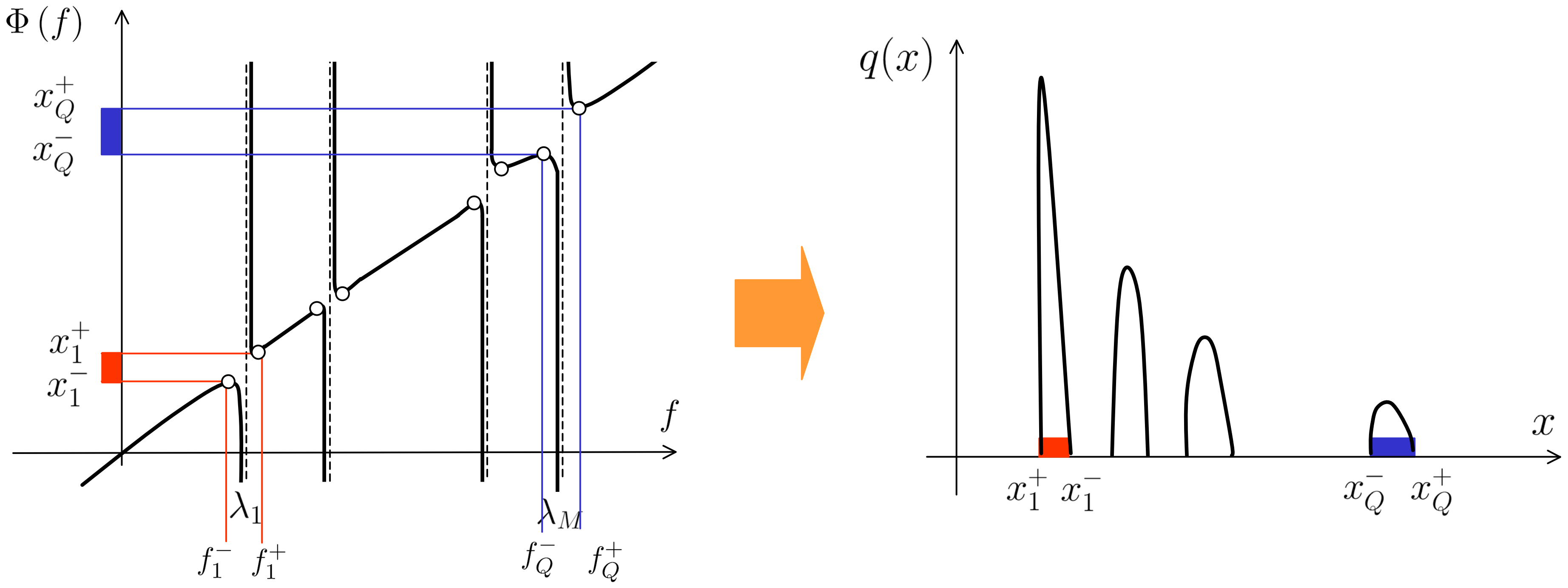
Hence, the support of $q(x)$ can be expressed as $\mathcal{S} = \left[x_1^-, x_1^+\right] \cup \ldots \cup \left[x_Q^-, x_Q^+\right]$.

Note: It may happen that $\lambda_m \notin \left(x_q^-, x_q^+\right)$. even when $\lambda_m \in \left(f_q^-, f_q^+\right)$.

Determining the support of $q(x)$



Determining the support of $q(x)$

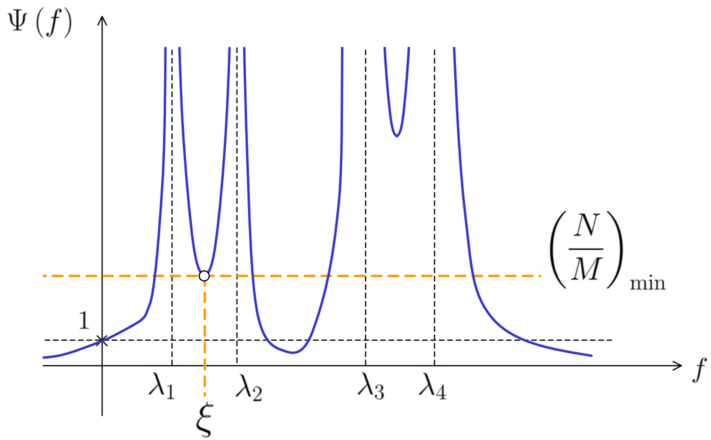


Splitting number for two clusters in $q(x)$

Given two consecutive eigenvalues $\{\lambda_m, \lambda_{m+1}\}$ there exists a minimum number of samples per antenna to guarantee that the corresponding eigenvalue clusters split.

$$\left(\frac{N}{M}\right)_{\min} > \frac{1}{M}\sum_{k=1}^{M}\left(\frac{\lambda_k}{\lambda_k - \xi}\right)^2$$

where ξ is the mth real valued solution to $\Psi'(f) = 0$.



Sketch of the derivation

The quantity that needs to be analyzed can be written as

$$\eta(\theta) = \mathbf{s}^H(\theta)\,\hat{\mathbf{E}}_N\hat{\mathbf{E}}_N^H\mathbf{s}(\theta) = \|\mathbf{s}(\theta)\|^2 - \mathbf{s}^H(\theta)\,\hat{\mathbf{E}}_S\hat{\mathbf{E}}_S^H\mathbf{s}(\theta)$$

$$\mathbf{s}^H(\theta)\,\hat{\mathbf{E}}_S\hat{\mathbf{E}}_S^H\mathbf{s}(\theta) = \frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{C}^-}\hat{m}_M(z)dz$$

where

$$\hat{m}_M(z) = \mathbf{s}^H(\theta)\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\mathbf{s}(\theta)$$

and $\mathcal{C}^-$ is a negatively oriented contour enclosing only the signal sample eigenvalues $\left\{\hat{\lambda}_{M-K+1},\ldots,\hat{\lambda}_M\right\}$.

The idea of the derivation is based on obtaining a proper parametrization of the contour $\mathcal{C}^-$.

Sketch of the derivation (ii)

We can therefore apply classical convergence results ($|\hat{m}_M(z) - \bar{m}_M(z)| \to 0$) to obtain

$$\left| \frac{1}{2\pi \mathrm{j}} \oint_{\mathcal{C}^-} \hat{m}_M(z) dz - \frac{1}{2\pi \mathrm{j}} \oint_{\mathcal{C}^-} \bar{m}_M(z) dz \right| \to 0$$

where, we recall that

$$\bar{m}_M(z) = \frac{f_M(z)}{z} \mathbf{a}^H(\theta) \left(\mathbf{R}_M - f_M(z) \mathbf{I}_M \right)^{-1} \mathbf{a}(\theta).$$

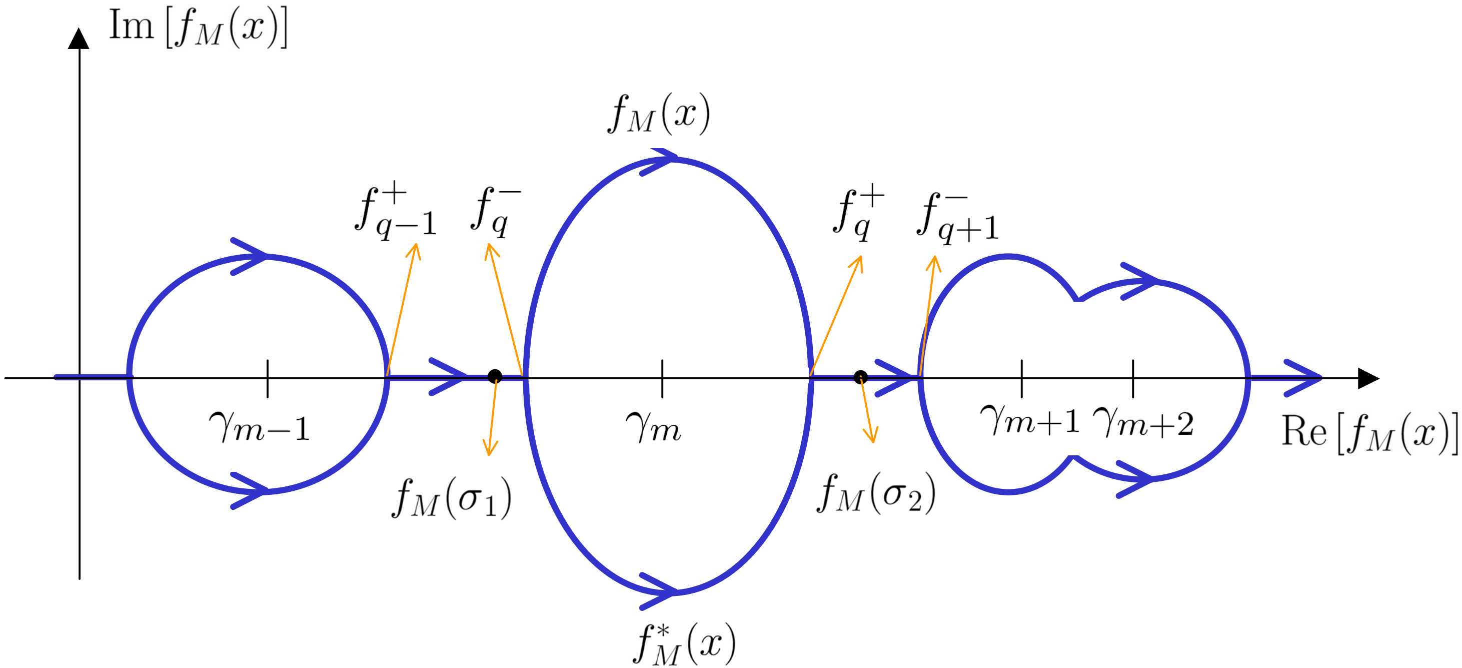
and where $f_M(z)$ is related to the Stieltjes transform of the eigenvalues, and is found as the solution in $\mathbb{C}^+$ of

$$z = f \left(1 - \frac{c}{M} \sum_{m=1}^{M} \frac{\lambda_m}{\lambda_m - f} \right)$$

We are implicitely using localization properties of the sample eigenvalues.

Sketch of the derivation (iii)

It turns out that $f_M(x)$ gives us a valid parametrization of $\mathcal{C}^-$:



Sketch of the derivation (iv)

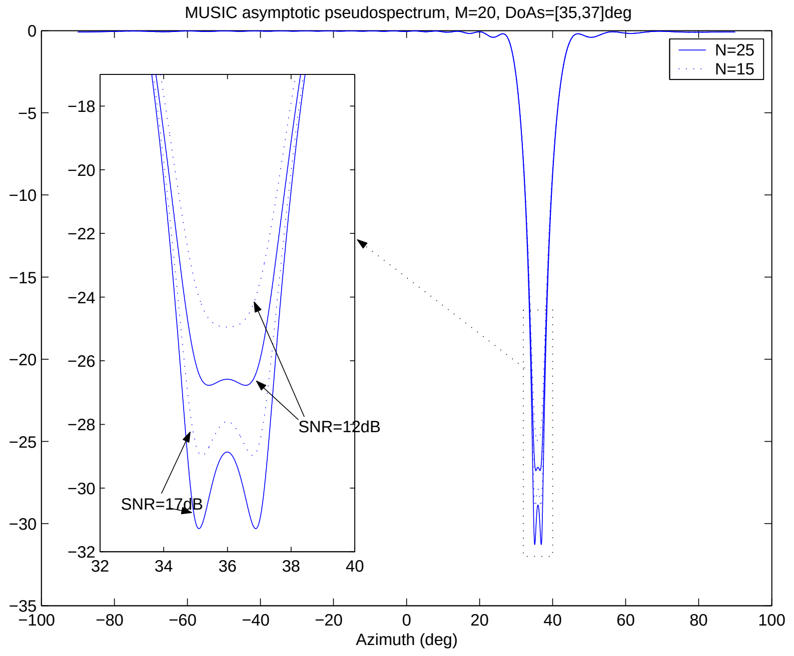
Using this fact in the asymptotic expression of the MUSIC estimator, we obtain

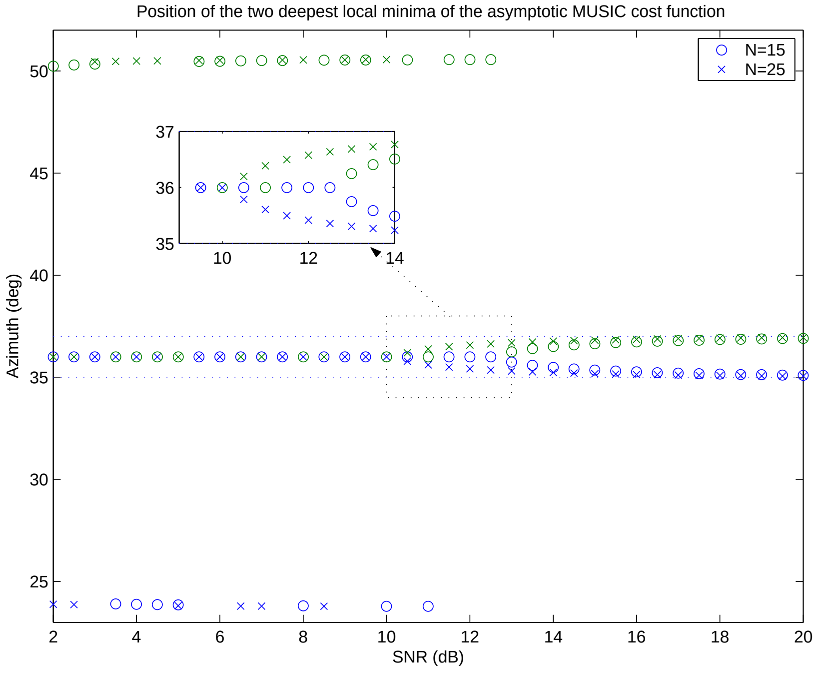
$$
\begin{aligned}
\frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{C}^-}\bar{m}_M(z)dz &= \frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{C}^-}\mathbf{a}^H(\theta)\left(\mathbf{R}_M - f_M(z)\mathbf{I}_M\right)^{-1}\mathbf{a}(\theta)\frac{f_M(z)}{z}dz \\
&= \frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{C}^-}\mathbf{a}^H(\theta)\left(\mathbf{R}_M - f_M(z)\mathbf{I}_M\right)^{-1}\mathbf{a}(\theta)\frac{1-\frac{c}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m - f_M(z)}\right)^2}{1-\frac{c}{M}\sum_{m=1}^{M}\frac{\lambda_m}{\lambda_m - f_M(z)}}f_M'(z)dz \\
&= \sum_{k=1}^{M}\mathbf{a}^H(\theta)\mathbf{e}_k\mathbf{e}_k^H\mathbf{a}(\theta)\frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{R}^-}\frac{1-\frac{c}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m-\omega}\right)^2}{(\lambda_k-\omega)\left(1-\frac{c}{M}\sum_{m=1}^{M}\frac{\lambda_m}{\lambda_m-\omega}\right)}d\omega
\end{aligned}
$$

from which the asymptotic formula for the MUSIC estimator follows.

Asymptotic behavior of MUSIC: an example

We consider a scenario with two sources impinging on a ULA ($d/\lambda_c = 0.5$, $M = 20$) from DoAs: $35°$, $37°$.





Estimation under low sample support with random matrix theory

When designing an estimator of a certain scalar function of $\mathbf{R}_M$, namely $\varphi(\mathbf{R}_M)$, one can distinguish between:

- **Traditional** N**-consistency:** Consistency when $N \to \infty$ while M remains fixed.
- M, N**-consistency:** Consistency when $M, N \to \infty$ at the same rate.

We observe that M, N-consistency guarantees a good behavior when the number of samples N has the same order of magnitude as the observation dimension M.

The objective of G-estimation (V.L. Girko) is to provide a systematic approach for the derivation of M, N-consistent estimators of different scalar functions of the true covariance matrix. For example, the G-estimator of $\frac{1}{M} \operatorname{tr}\left[\mathbf{R}_M^{-1}\right]$ will be

$$\frac{(1-c)}{M} \operatorname{tr}\left[\hat{\mathbf{R}}_M^{-1}\right].$$

M, N-consistent subspace detection: G-MUSIC

We propose to use an M, N-consistent estimator (under separability conditions) of the cost function $\eta(\theta) = \mathbf{s}^H(\theta)\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\theta)$:

$$\eta_{\text{G-MUSIC}}(\theta) = \mathbf{s}^H(\theta)\left(\sum_{k=1}^{M}\phi(k)\hat{\mathbf{e}}_k\hat{\mathbf{e}}_k^H\right)\mathbf{s}(\theta)$$

$$\phi(k) = \begin{cases} 1 + \sum_{r=M-K+1}^{M}\left(\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r}\right) & k \leq M-K \\ -\sum_{r=1}^{M-K}\left(\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r}\right) & k > M-K \end{cases}$$

where now $\hat{\mu}_1, \ldots, \hat{\mu}_M$ are the solutions to the equation

$$\frac{1}{M}\sum_{k=1}^{M}\frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\mu}} = \frac{1}{c}.$$

Sketch of the derivation

The quantity that needs to be estimated can be written as

$$\eta(\theta) = \mathbf{s}^H(\theta)\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\theta) = \|\mathbf{s}(\theta)\|^2 - \mathbf{s}^H(\theta)\mathbf{E}_S\mathbf{E}_S^H\mathbf{s}(\theta)$$

$$\mathbf{s}^H(\theta)\mathbf{E}_S\mathbf{E}_S^H\mathbf{s}(\theta) = \frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{C}^-}\mathbf{s}^H(\theta)\left(\mathbf{R}_M - \omega\mathbf{I}_M\right)^{-1}\mathbf{s}(\theta)\,d\omega$$

where $\mathcal{C}^-$ is a negatively oriented contour enclosing only the signal eigenvalues $\{\lambda_{M-K+1}, \ldots, \lambda_M\}$.

The idea of the derivation is based on using, once again, $f_M(x)$ as a parametrization of the contour $\mathcal{C}^-$.

$$\eta(\theta) = \mathbf{s}^H(\theta)\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\theta) = \|\mathbf{s}(\theta)\|^2 - \frac{1}{\pi}\operatorname{Im}\left[\int_{\sigma_1}^{\sigma_2}\mathbf{s}^H(\theta)\left(\mathbf{R}_M - f_M(x)\mathbf{I}_M\right)^{-1}\mathbf{s}(\theta)\,f_M'(x)dx\right]$$

Sketch of the derivation (ii)

From this point, we can express the integral in terms of $\bar{m}_M(z)$ and $\bar{b}_M(z)$, namely

$$\eta\left(\theta\right) = \left\|\mathbf{s}\left(\theta\right)\right\|^2 - \frac{1}{\pi}\operatorname{Im}\left[\int_{\sigma_1}^{\sigma_2} \bar{m}_M(x)\frac{1-c+cx^2\bar{b}'_M(x)}{1-c-cx\bar{b}_M(x)}dx\right]$$

and we only need to replace $\bar{m}(z)$ and $\bar{b}(z)$ with their M, N-consistent estimates:

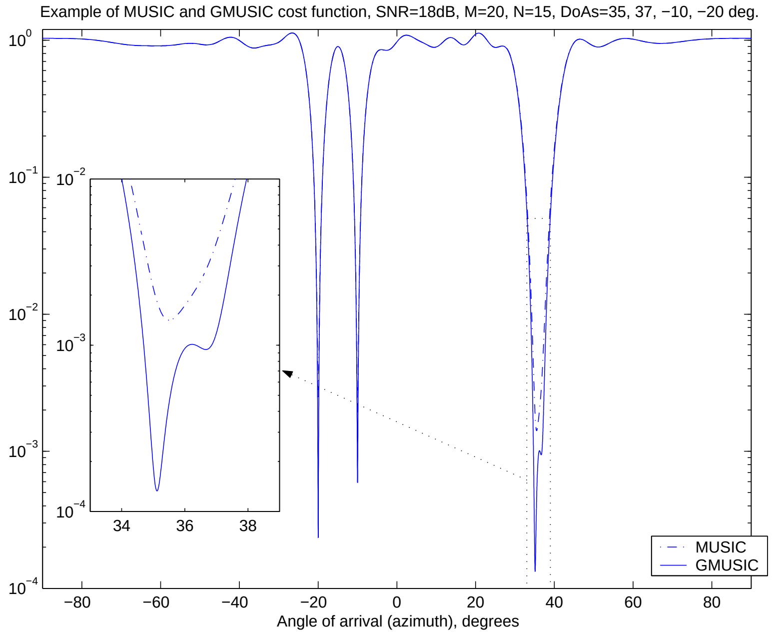
$$\eta_{\textsf{G-MUSIC}}\left(\theta\right) = \left\|\mathbf{s}\left(\theta\right)\right\|^2 - \frac{1}{\pi}\operatorname{Im}\left[\int_{\sigma_1}^{\sigma_2} \hat{m}_M(x)\frac{1-c+cx^2\hat{b}'_M(x)}{1-c-cx\hat{b}_M(x)}dx\right].$$

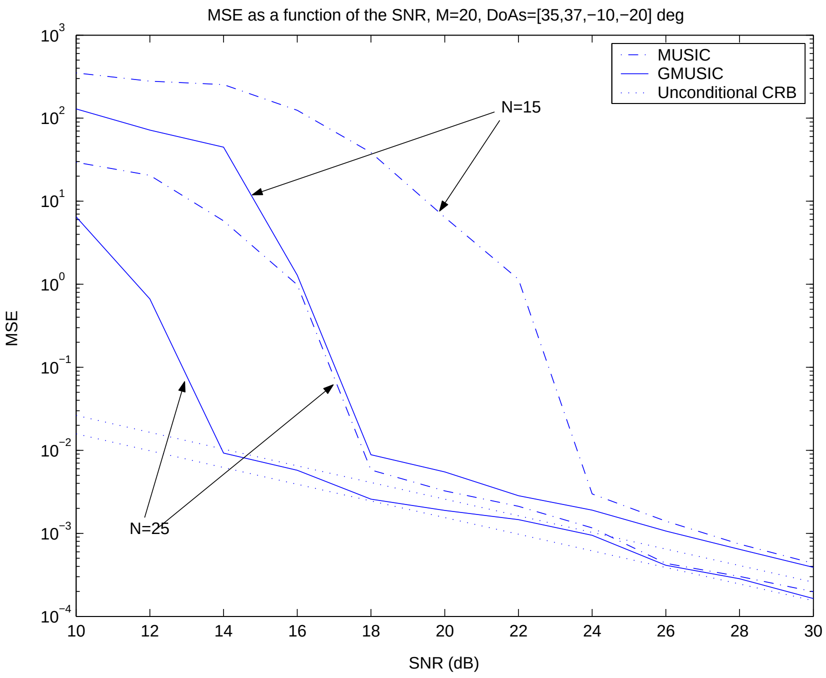
Integrating this expression, we get to the proposed estimator:

$$\eta_{\textsf{G-MUSIC}}\left(\theta\right) = \mathbf{s}^H\left(\theta\right)\left(\sum_{k=1}^{M}\phi(k)\hat{\mathbf{e}}_k\hat{\mathbf{e}}_k^H\right)\mathbf{s}\left(\theta\right).$$

Performance evaluation MUSIC vs. G-MUSIC

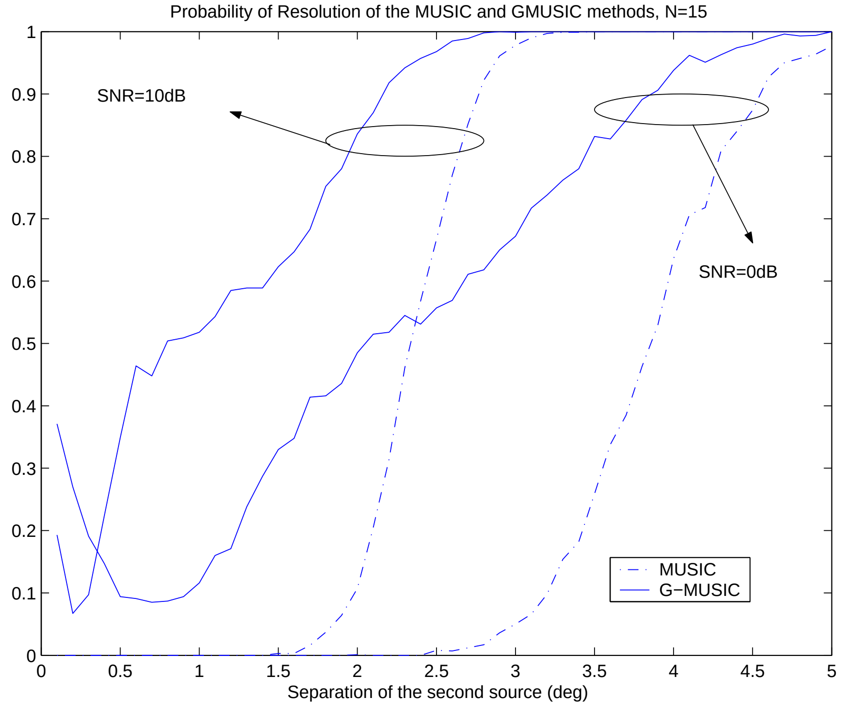
Comparative evaluation of MUSIC and G-MUSIC via simulations. Scenario with four sources ($35°$, $37°$, $-10°$, $-20°$), ULA ($M = 20$, $d/\lambda_c = 0.5$).

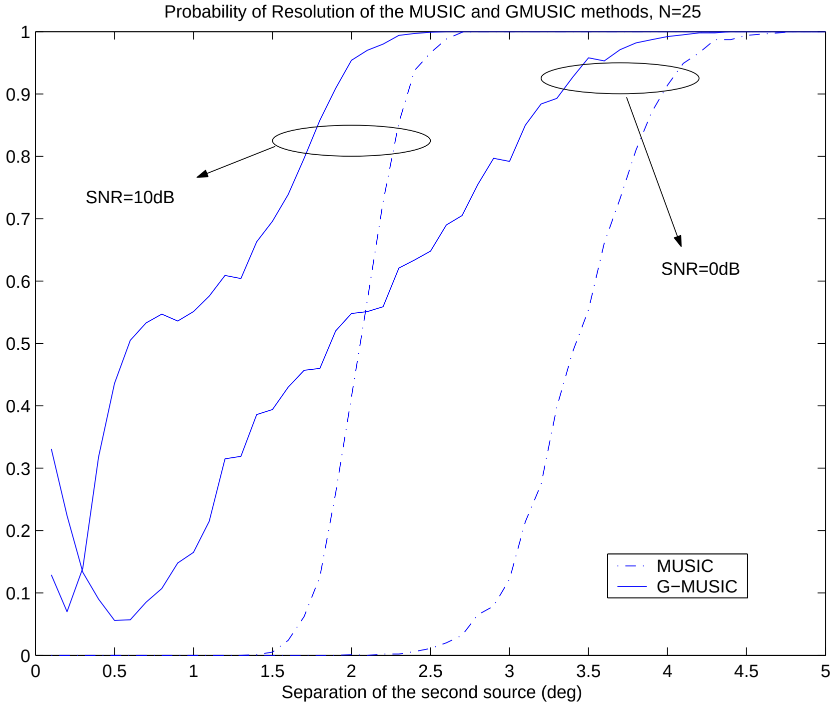




Performance evaluation MUSIC vs. G-MUSIC (ii)

Resolution capabilities of MUSIC and G-MUSIC via simulations. Scenario with two sources ($\theta_1 = 0°$, $\theta_2 = 0° \ldots 5°$), ULA ($M = 20$, $d/\lambda_c = 0.5$). $\text{Prob}\left(\eta\left(\frac{\theta_1+\theta_2}{2}\right) > \max\left\{\eta\left(\theta_1\right), \eta\left(\theta_2\right)\right\}\right)$.





Power estimation

- Once the DoAs of the K different sources impinging on the array have been estimated, it remains the problem of determining their corresponding power $P_1, \ldots, P_K$. We assume that sources are uncorrelated.
- A very common approach is to use the Gaussian Maximum Likelihood estimator obtained without imposing zero correlation between the different sources, namely

$$\hat{P}_k^{\mathsf{ML}} = \left\{ \mathbf{A}^{\#}(\hat{\Theta}) \left(\hat{\mathbf{R}}_M - \hat{\sigma}^2_{\mathsf{ML}} \mathbf{I}_M \right) \mathbf{A}^{\#}(\hat{\Theta})^H \right\}_{kk}$$

where $\mathbf{A}^{\#}(\hat{\Theta}) = \left(\mathbf{A}^H(\hat{\Theta}) \mathbf{A}(\hat{\Theta}) \right)^{-1} \mathbf{A}(\hat{\Theta})^H$, and

$$\hat{\sigma}^2_{\mathsf{ML}} = \frac{1}{M-K} \operatorname{tr} \left[\hat{\mathbf{R}} \mathbf{P}_A^{\perp} \right], \quad \mathbf{P}_A^{\perp} = \mathbf{I}_M - \mathbf{A}(\hat{\Theta}) \mathbf{A}^{\#}(\hat{\Theta}).$$

- The power estimate associated with a particular source depends on the DoA of all the sources in the scenario. Therefore, the presence of one outlier in the DoA estimation may adversely effect even the power estimation of the sources whose DoA has been correctly detected.

Power estimation depending on the desired DoA

We will consider the estimator proposed in [McCloud Scharf 2002], based on the following identity

$$P_k = \frac{1}{\mathbf{s}^H(\theta_k)\left(\mathbf{R}-\sigma^2\mathbf{I}_M\right)^{\#}\mathbf{s}(\theta_k)}$$

where $(\cdot)^{\#}$ denotes the Moore-Penrose pseudoinverse. Based on this identity in, and noting that we can express

$$\left(\mathbf{R}-\sigma^2\mathbf{I}_M\right)^{\#} = \mathbf{E}_S\left(\mathbf{\Lambda}_S-\sigma^2\mathbf{I}_K\right)^{-1}\mathbf{E}_S^H$$

the following estimator of the signal power of the kth source was proposed:

$$\hat{P}_k = \frac{1}{\mathbf{s}^H(\theta_k)\hat{\mathbf{E}}_S\left(\hat{\mathbf{\Lambda}}_S-\hat{\sigma}^2\mathbf{I}_K\right)^{-1}\hat{\mathbf{E}}_S^H\mathbf{s}(\theta_k)}$$

where $\hat{\sigma}^2 = \frac{1}{M-K}\sum_{k=1}^{M-K}\hat{\lambda}_k$ is an estimator of the noise power.

Power estimation depending on the desired DoA (ii)

- Contrary to the Maximum Likelihood solution, the proposed estimator depends only on the signature of the source of interest (it can be used in reduced angular explorations and it is robust to the presence of outliers in the DoA detection of the other sources).
- Furthermore, it can be implemented efficiently once the eigendecomposition of the sample covariance matrix is available from a previous stage:

$$\hat{P}_k = \frac{1}{\sum_{m=M-K+1}^{M} \frac{1}{\hat{\lambda}_m - \hat{\sigma}^2} \left| \mathbf{s}^H(\theta_k) \hat{\mathbf{e}}_m \right|^2}$$

- The main drawback comes from the fact that it has been obtained by direct replacement of $\mathbf{R}$ by $\hat{\mathbf{R}}$, which only ensures N-consistency. Consequently, the performance when M, N have a comparable order of magnitude, the estimator may give quite biased estimates.
- A potential way of circumventing this problem consists in designing an M,N-consistent estimator of the original quantity.

M, N-consistent power estimation depending on the desired DoA

Under some assumptions, the following estimator of the power of the kth source is strongly consistent as $M, N \to \infty$ at the same rate,

$$\check{P}_k = \frac{1}{\mathbf{s}^H(\theta_k)\left(\sum_{m=1}^{M} \psi(m)\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H\right)\mathbf{s}(\theta_k)}$$

where $\psi(m)$, $m = 1 \dots M$ are the coefficients

$$\psi(m) = \begin{cases} -\frac{1}{\check{\sigma}^2}\sum_{k=M-K+1}^{M}\left(\frac{\hat{\vartheta}_k}{\hat{\lambda}_m - \hat{\vartheta}_k} - \frac{\hat{\lambda}_k}{\hat{\lambda}_m - \hat{\lambda}_k}\right) & m \leq M-K \\ \frac{1}{\check{\sigma}^2}\left(\sum_{k=0}^{M-K}\frac{\hat{\vartheta}_k}{\hat{\lambda}_m - \hat{\vartheta}_k} - \sum_{k=1}^{M-K}\frac{\hat{\lambda}_k}{\hat{\lambda}_m - \hat{\lambda}_k}\right) & m > M-K. \end{cases}$$

Even though we need to estimate something from the signal subspace, the estimator combines information from both subspaces in the sample covariance matrix.

M, N-consistent power estimation depending on the desired DoA (ii)

In the last expression, $\hat{\vartheta}_k$ are the real-valued solutions to the following equations in $\hat{\mu}$ and $\hat{\vartheta}$ respectively

$$\hat{\vartheta} = \breve{\sigma}^2 \left(1 - c \frac{1}{M} \sum_{k=1}^{M} \frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\vartheta}} \right).$$

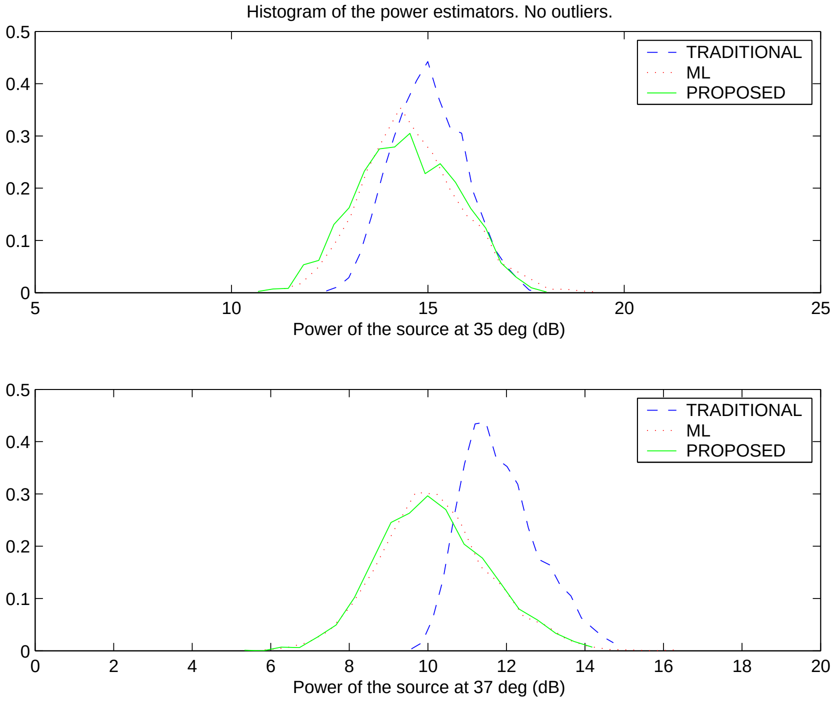
and $\breve{\sigma}^2$ is an estimator of the noise power σ^2 that is also strongly consistent as $M, N \to \infty$ at the same rate

$$\breve{\sigma}^2 = \frac{N}{M-K} \sum_{k=1}^{M-K} \left(\hat{\lambda}_k - \hat{\mu}_k \right), \qquad \frac{1}{M} \sum_{k=1}^{M} \frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\mu}} = \frac{1}{c}$$

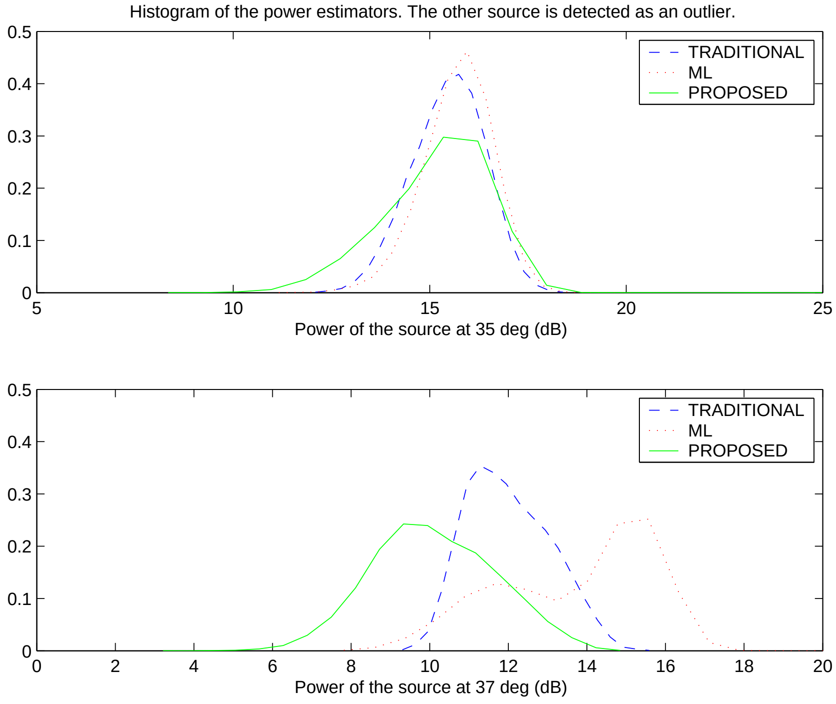
It can be shown that, as $c \to 0$ (increasingly high number of samples per antenna), the estimator reverts to the classical one.

Numerical evaluation

- Scenario $K = 2$, ULA $M = 20$. DoAs of 35 and 37 degrees with a power 15dB and 10dB above the noise floor respectively. $N = 35$. GMUSIC + Power estimation.



(a) No outliers



(b) The other source DoA is an outlier estimation

Conclusions

- Random Matrix Theory is a very useful framework to characterize and model the finite sample size regime in array signal processing applications.
- In particular, we have seen that we can provide closed form expressions for the asymptotic pseudospectrum of the MUSIC algorithm when both the number of antennas and the sample volume increase at the same rate.
- Using this framework, we can propose estimators that are consistent, not only when the number of snapshots tends to infinity ($N \to \infty$), but also when the number of antennas increases as well.
- We have derived M, N-consistent estimators of the MUSIC cost function and the corresponding power estimators for each of the sources.

Some references

- Xavier Mestre, "On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices", IEEE Transactions on Signal Processing, vol. 56, no. 11, November 2008.
- Xavier Mestre, "Improved estimation of eigenvalues of covariance matrices and their associated subspaces using their sample estimates," IEEE Transactions on Information Theory, vol. 54, no. 11, Nov. 2008.
- Xavier Mestre, Miguel A. Lagunas, "Modified subspace algorithms for DoA detection in the small sample size regime," IEEE Transactions on Signal Processing, 2007.
- Xavier Mestre, Ben A. Johnson, Yuri I. Abramovich, "Source power estimation for array processing applications under low sample size constraints", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2007.

Thank you for your attention!!!