Some Applications of Large Random Matrices to Array Processing

Ph. Loubaton

LIGM, Université Paris-Est Marne la Vallée.

Summer School, 7 June 2016

1. Problem statement

2. The narrow band array processing model
   - Detailed presentation of the narrow band array processing model.
   - The pure noise case: the Marcenko-Pastur distribution
   - The signal plus noise case.
   - Applications.

3. Wideband array processing models.
   - Detailed description of the wideband array processing model
   - Asymptotic behaviour of the empirical spatio-temporal covariance matrix.
   - Applications.

4. Conclusion

The random matrix models considered in this lecture I.

I. The narrow band array processing model.

- $M \times N$ random matrices, M number of sensors, N number of snapshots
- Random matrix model $\mathbf{Y} = \mathbf{AS} + \mathbf{V}$
- $\mathbf{V}$ complex Gaussian i.i.d. random matrix modelling the additive noise
- $\mathbf{A}$ the $M \times K$ matrix of "directional vectors", $K << M$ number of sources
- $\mathbf{S}$ the $K \times N$ deterministic matrix collecting the source signals

When M and N are large and K small:

- Behaviour of the largest eigenvalues and associated eigenvectors of $\frac{\mathbf{YY}^*}{N}$
- Detection: testing $K = 0$ versus $K = K_0$
- Dimension reduction by PCA
- Behaviour of subspace DoA estimators

II. Wide band or spatio-temporal array processing model.

- $ML \times N$ block Hankel random matrices, M number of sensors, N number of snapshots, L a smoothing factor
- $\mathbf{Y}^{(L)} = (\mathbf{Y}_1^{(L)T}, \ldots, \mathbf{Y}_M^{(L)T})^T$
- Each block $\mathbf{Y}_k^{(L)}$ is a $L \times N$ Hankel matrix built from the signal $(y_k(n))_{n=1,\ldots,N}$ observed on sensor k
- $\mathbf{Y}^{(L)} = \mathbf{H}^{(L)}\mathbf{S}^{(L)} + \mathbf{V}^{(L)}$
- $\mathbf{V}^{(L)}$ is this time a Gaussian random block Hankel matrix
- The signal part $\mathbf{H}^{(L)}\mathbf{S}^{(L)}$ is a low rank K deterministic matrix

When M, L, N are large and K small:

- Behaviour of the largest eigenvalues and eigenvectors of $\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}}{N}$
- Application to detection of a wideband signal
- Loading factor estimation for trained regularized spatio-temporal Wiener filtering
- Analysis of spatial smoothing schemes in narrow band array processing

1. Problem statement

2. The narrow band array processing model
   - Detailed presentation of the narrow band array processing model.
   - The pure noise case: the Marcenko-Pastur distribution
   - The signal plus noise case.
   - Applications.

3. Wideband array processing models.
   - Detailed description of the wideband array processing model
   - Asymptotic behaviour of the empirical spatio-temporal covariance matrix.
   - Applications.

4. Conclusion

The narrow band array processing model

Observation: M–dimensional time series $\mathbf{y}_n$ observed from $n = 1, \ldots, N$

- $\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n$
- $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_K)$ deterministic unknown rank $K < M$ matrix
- $\mathbf{s}_n = (s_{1,n}, \ldots, s_{K,n})^T$, $((s_{k,n})_{n \in \mathbb{Z}})_{k=1,K}$ are $K < M$ non observable deterministic "source signals"
- $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ additive complex white Gaussian noise such that $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^H) = \sigma^2 \mathbf{I}_M$

In matrix form

- $\mathbf{Y}_N = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$ observation $M \times N$ matrix
- $\mathbf{S}_N = (\mathbf{s}_1, \ldots, \mathbf{s}_N)$ signal $K \times N$ matrix, $\mathrm{Rank}(\mathbf{S}_N) = K$.
- $\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$ Information + Noise model with rank deficient Information component.

The narrow band array processing model

In matrix form

- $\mathbf{Y}_N = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$ observation $M \times N$ matrix
- $\mathbf{S}_N = (\mathbf{s}_1, \ldots, \mathbf{s}_N)$ signal $K \times N$ matrix, $\mathrm{Rank}(\mathbf{S}_N) = K$.
- $\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$ Information + Noise model with rank deficient Information component.

The asymptotic regime.

- Easier to design and study statistical inference methods in asymptotic regimes.
- If $M << N$: M fixed and $N \to +\infty$
- If M and N are of the same order of magnitude: $M \to +\infty, N \to +\infty$ in such a way that $c_N = \frac{M}{N} \to c_*$, $0 < c_* < +\infty$

Class of problems to be addressed.

Covariance matrices of the model.

- $\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$
- Empirical covariance matrix $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n \mathbf{y}_n^*$
- "True" covariance matrix $\mathbb{E}\left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}\right) = \mathbf{A}\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}\mathbf{A}^* + \sigma^2 \mathbf{I}_M$

Extract informations on $\frac{\mathbf{A}\mathbf{S}_N \mathbf{S}_N^* \mathbf{A}^*}{N}$ from $\mathbf{Y}_N$.

- If M fixed and $N \rightarrow +\infty$, classical problems because

$$\|\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} - \left(\mathbf{A}\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}\mathbf{A}^* + \sigma^2 \mathbf{I}_M\right)\| \rightarrow 0$$

- If $M \rightarrow +\infty, N \rightarrow +\infty$ in such a way that $c_N = \frac{M}{N} \rightarrow c_*$, $0 < c_* < +\infty$, this property does not hold

Class of problems to be addressed.

Covariance matrices of the model.

- $\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$
- Empirical covariance matrix $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n \mathbf{y}_n^*$
- "True" covariance matrix $\mathbb{E}\left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}\right) = \mathbf{A} \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \mathbf{A}^* + \sigma^2 \mathbf{I}_M$

Extract informations on $\frac{\mathbf{A}\mathbf{S}_N \mathbf{S}_N^* \mathbf{A}^*}{N}$ from $\mathbf{Y}_N$ in the asymptotic regime.

- $M = M(N)$, $N \to +\infty$ in such a way that $c_N = \frac{M(N)}{N} \to c_*$, $0 < c_* < 1$
- Written as $N \to +\infty$
- K does not scale with (M, N)

In some sense, M depends on N. We denote $M \times K$ matrix $\mathbf{A}$ by $\mathbf{A}_N$.

Properties of the empirical covariance matrix when $\mathbf{Y} = \mathbf{V}$ $(K = 0)$.

$$\mathbf{V} = \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1N} \\ V_{21} & V_{22} & \dots & V_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ V_{M1} & V_{M2} & \dots & V_{MN} \end{pmatrix}$$

$(V_{ij})_{1 \leq i \leq M, 1 \leq j \leq N}$ i.i.d. complex Gaussian random variables $\mathcal{CN}(0, \sigma^2)$.
$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ columns of $\mathbf{V}$, $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}_M$

Empirical covariance matrix:

$$\frac{\mathbf{V}\mathbf{V}^*}{N} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n^*$$

Behaviour of the empirical distribution of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$ for large M and N.

- $\hat{\lambda}_{1,N} \geq \hat{\lambda}_{2,N} \geq \ldots \geq \hat{\lambda}_{M,N}$ eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$
- Empirical eigenvalue distribution: $\hat{\mu}_N = \frac{1}{M}\sum_{i=1}^{M} \delta(\lambda - \hat{\lambda}_{i,N})$

How behave the histograms of the eigenvalues $(\hat{\lambda}_{i,N})_{i=1,\ldots,M}$ of $\frac{\mathbf{V}\mathbf{V}^*}{N}$ when M and N increase.

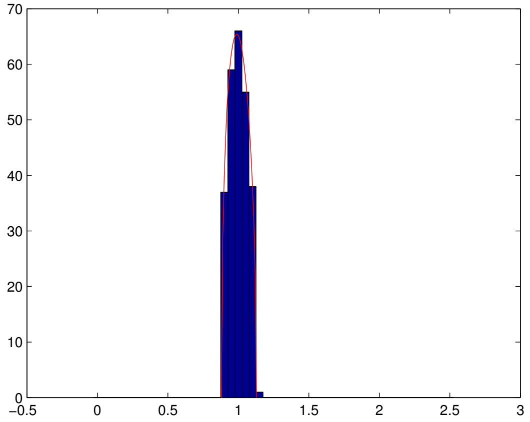
Well known case: M fixed, N increases i.e. $c_N = \frac{M}{N}$ small

$\frac{\mathbf{V}\mathbf{V}^*}{N} \simeq \mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}_M$ by the law of large numbers.

If $N >> M$, the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$ are concentrated around σ^2.

Illustration.

Histogram of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$, $M = 256, c_N = \frac{M}{N} = \frac{1}{256}$, $\sigma^2 = 1$



If M et N are of the same order of magnitude.

- $\left(\frac{\mathbf{V}\mathbf{V}^*}{N}\right)_{i,j} \to 0$ if $i \neq j$
- $\left(\frac{\mathbf{V}\mathbf{V}^*}{N}\right)_{i,j} \to \sigma^2$ if $i = j$
- But $\|\frac{\mathbf{V}\mathbf{V}^*}{N} - \sigma^2 \mathbf{I}_M\|$ does not converge towards 0.

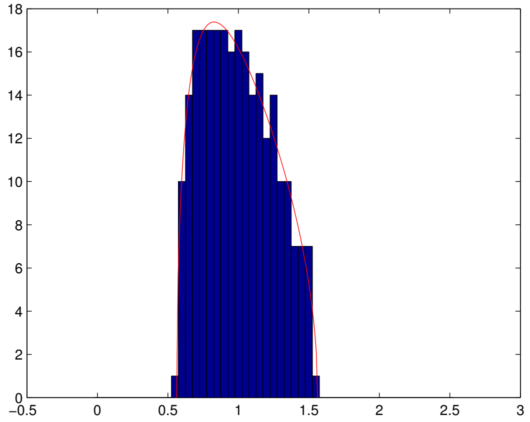
The histograms of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$ tend to concentrate around the probability density of the so-called Marcenko-Pastur distribution $\mathrm{MP}(\sigma^2, c_N)$: if $c_N \leq 1$

$$
\begin{aligned}
p_{\sigma^2, c_N}(\lambda) &= \frac{1}{2\pi c_N \lambda} \sqrt{[\sigma^2(1+\sqrt{c_N})^2 - \lambda][\lambda - \sigma^2(1-\sqrt{c_N})^2]} \\
& \quad \text{if } \lambda \in [\sigma^2(1-\sqrt{c_N})^2, \sigma^2(1+\sqrt{c_N})^2] \\
&= 0 \ \text{ otherwise}
\end{aligned}
$$

Result still true in the non Gaussian case

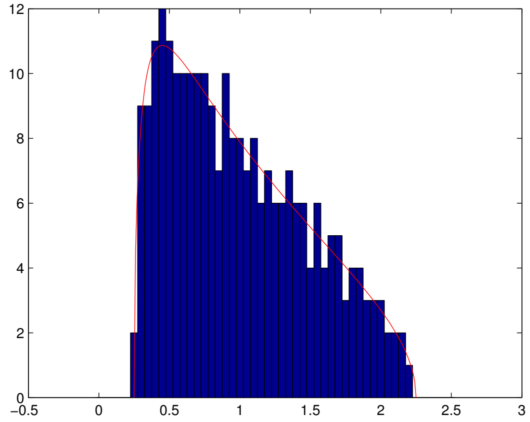
Illustrations I.

Histogram of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$, $M = 256, c_N = \frac{M}{N} = \frac{1}{16}$, $\sigma^2 = 1$



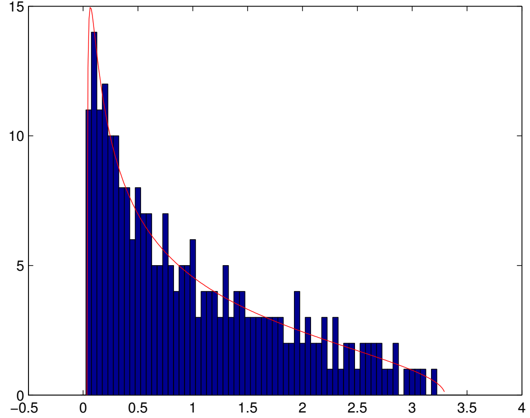
Illustrations II.

Histogram of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$, $M = 256, c_N = \frac{M}{N} = \frac{1}{4}$, $\sigma^2 = 1$



Illustrations III.

Histogram of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$, $M = 256, c_N = \frac{M}{N} = 2/3$, $\sigma^2 = 1$



More formally

$$\frac{1}{M}\sum_{k=i}^{M}\psi(\hat{\lambda}_{i,N}) - \int \psi(\lambda)p_{\sigma^2,c_N}(\lambda)\,d\lambda \to 0$$

Fluctuations of the linear statistics of the $(\hat{\lambda}_{i,N})_{i=1,\ldots,M}$.

- $\mathrm{Var}\left(\frac{1}{M}\sum_{k=i}^{M}\psi(\hat{\lambda}_{i,N})\right) = \mathcal{O}(\frac{1}{N^2})$
- $\mathbb{E}\left(\frac{1}{M}\sum_{i=1}^{M}\psi(\hat{\lambda}_{i,N})\right) - \int \psi(\lambda)p_{\sigma^2,c_N}(\lambda)\,d\lambda \ = \mathcal{O}(\frac{1}{N^2})$
- $N\left[\left(\frac{1}{M}\sum_{i=1}^{M}\psi(\hat{\lambda}_{i,N})\right) - \int \psi(\lambda)p_{\sigma^2,c_N}(\lambda)\,d\lambda\right] \to \mathcal{N}(0,\Delta)$

The $(\hat{\lambda}_{i,N})_{i=1,\ldots,M}$ do not behave at all as realizations of independent random variables.

Finer convergence results.

Convergence of the extreme eigenvalues

$$\hat{\lambda}_{1,N} - \sigma^2(1+\sqrt{c_N})^2 \xrightarrow[N,M\to\infty]{a.s.} 0$$
$$\hat{\lambda}_{M,N} - \sigma^2(1-\sqrt{c_N})^2 \xrightarrow[N,M\to\infty]{a.s.} 0$$

Implies the following almost sure location property of the $(\hat{\lambda}_{i,N})_{i=1,\ldots,M}$.

- For each $\epsilon > 0$, almost surely, all the eigenvalues belong to $[\sigma^2(1-\sqrt{c_N})^2 - \epsilon, \sigma^2(1+\sqrt{c_N})^2 + \epsilon]$ for N large enough.
- Important property valid in the context of other models based on i.i.d. complex Gaussian matrices (Bai-Silverstein 1999 for the zero mean correlated case, Haagerup 2005, Male 2012).

Fluctuations of the extreme eigenvalues.

A Central Limit Theorem holds for the largest eigenvalue $\hat{\lambda}_{1,N}$. When correctly centered and rescaled, $\hat{\lambda}_{1,N}$ converges to a **Tracy-Widom** distribution:

$$\frac{N^{2/3}}{\sigma^2} \times \frac{\hat{\lambda}_{1,N} - \sigma^2(1+\sqrt{c_N})^2}{(1+\sqrt{c_N})\left(\frac{1}{\sqrt{c_N}}+1\right)^{1/3}} \xrightarrow[N,M\to\infty]{\mathcal{L}} \mu_{TW} \ .$$

The function μ_{TW} stands for **Tracy-Widom** distribution.

A similar result holds for $\hat{\lambda}_{M,N}$, the smallest eigenvalue.

The model.

We recall that:

$$\begin{array}{ccccccc} \text{Rcv signal} & & \text{Channel} & \text{Src signal} & & \text{Noise} \\ \begin{bmatrix} \mathbf{y}_1 \cdots \mathbf{y}_N \end{bmatrix} & = & \begin{bmatrix} \mathbf{a}_{N,1} \cdots \mathbf{a}_{N,K} \end{bmatrix} & \begin{bmatrix} \mathbf{s}^1 \\ \cdots \\ \mathbf{s}^K \end{bmatrix} & + & \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_N \end{bmatrix} \\ \mathbf{Y}_N & = & \mathbf{A}_N & \mathbf{S}_N & + & \mathbf{V}_N \\ M \times N & & M \times K & K \times N & & M \times N \end{array}$$

Asymptotic regime: $N \to \infty$, $c_N = M/N \to c_*$, and K is fixed.

$\mathbf{Y}_N$ = Matrix with Gaussian iid elements + fixed rank perturbation.

Results to be used when **number of sources** K **is** $\ll M$.

Normalizations of the signal contributions.

For each $k = 1, \ldots, K$.

- $\sup_N \frac{1}{N} \sum_{n=1}^{N} |s_{k,n}|^2 < +\infty$
- $\sup_N \|\mathbf{a}_{N,k}\|^2 < +\infty$

If the components of $\mathbf{a}_{N,k}$ are of the same order of magnitude $\mathcal{O}(\frac{1}{\sqrt{M}})$:

- SNR per sensor is $\mathcal{O}(\frac{1}{M}) \rightarrow 0$
- SNR at the output of each matched filter $\mathbf{a}_{N,k}^* \mathbf{y}_n$ is $\mathcal{O}(1)$

Notations

Spectral factorizations:

$$\frac{\mathbf{A}_N \mathbf{S}_N \mathbf{S}_N^* \mathbf{A}_N^*}{N} = \begin{bmatrix} \mathbf{u}_{1,N} & \cdots & \mathbf{u}_{K,N} \end{bmatrix} \begin{bmatrix} \lambda_{1,N} & & \\ & \ddots & \\ & & \lambda_{K,N} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{1,N} & \cdots & \mathbf{u}_{K,N} \end{bmatrix}^*$$

where $\lambda_{1,N} \geq \cdots \geq \lambda_{K,N}$.

$$\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} = \begin{bmatrix} \hat{\mathbf{u}}_{1,N} & \cdots & \hat{\mathbf{u}}_{M,N} \end{bmatrix} \begin{bmatrix} \hat{\lambda}_{1,N} & & \\ & \ddots & \\ & & \hat{\lambda}_{M,N} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_{1,N} & \cdots & \hat{\mathbf{u}}_{M,N} \end{bmatrix}^*$$

where $\hat{\lambda}_{1,N} \geq \cdots \geq \hat{\lambda}_{M,N}$.

Impact of the signal component on the eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$

If M is fixed and $N \to +\infty$, $c_N \simeq 0$

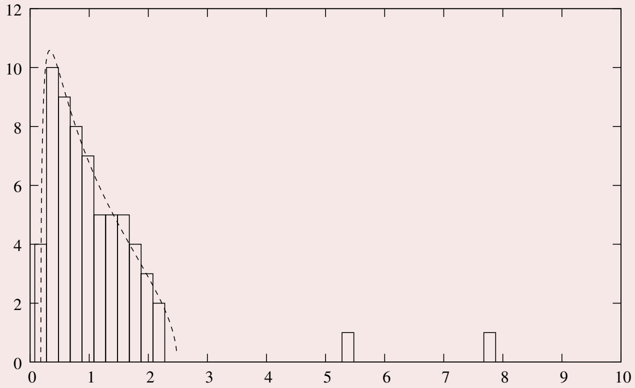
- $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} \simeq \mathbf{A}_N \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \mathbf{A}_N^* + \sigma^2 \mathbf{I}$
- $\hat{\lambda}_{k,N} \simeq \lambda_{k,N} + \sigma^2$ and $\hat{\mathbf{u}}_{k,N} \simeq \mathbf{u}_{k,N}$ if $1 \leq k \leq K$
- $\hat{\lambda}_{k,N} \simeq \sigma^2$ if $k > K$

In our asymptotic regime:

- The asymptotic distribution of $M - K$ smallest eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$ is the Marčenko Pastur
- Depending on the ratios $(\frac{\lambda_{k,N}}{\sigma^2})_{k=1,\ldots,K}$, at most K eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$ may escape from the support of the Marčenko Pastur and have a deterministic behaviour (more complicated than $\lambda_{k,N} + \sigma^2$)

Illustration

Histogram of the eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$, $c_N = \frac{M}{N} = 1/3$, $N = 192$, $K = 2$, $\lambda_1 = 6.25$, $\lambda_2 = 4$, $\sigma^2 = 1$



Main result on the eigenvalues

Theorem 1: Benaych-Georges and Nadakuditi, 2011

- Assume that $\lambda_{k,N} \to \rho_k$ for $k = 1, \ldots, K$.
- Let $i \leq K$ be the maximum index for which $\rho_i > \sigma^2 \sqrt{c_*}$ ($\lambda_{k,N} > \sigma^2 \sqrt{c_N}$ for $k \leq i$ and N large enough). Then for $k = 1, \ldots, i$,

$$\hat{\lambda}_{k,N} - \frac{\left(\sigma^2 c_N + \lambda_{k,N}\right)\left(\lambda_{k,N} + \sigma^2\right)}{\lambda_{k,N}} \xrightarrow[N\to\infty]{a.s.} 0$$

$$\gamma_{k,N} = \frac{\left(\sigma^2 c_N + \lambda_{k,N}\right)\left(\lambda_{k,N} + \sigma^2\right)}{\lambda_{k,N}} > \sigma^2 (1 + \sqrt{c_N})^2$$

and

$$\hat{\lambda}_{i+1,N} - \sigma^2 (1 + \sqrt{c_N})^2 \xrightarrow[N\to\infty]{a.s.} 0$$

Comments on Theorem I.

The almost sure location of the eigenvalues of $\frac{\mathbf{V}_N \mathbf{V}_N^*}{N}$ around the support of the MP distribution plays a fundamental role.

If $c_N \simeq 0$: $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} \simeq \mathbf{A}_N \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \mathbf{A}_N^* + \sigma^2 \mathbf{I}$ and $\hat{\lambda}_{k,N} \simeq (\lambda_{k,N} + \sigma^2)$.

$\lambda \rightarrow \frac{(\sigma^2 c_N + \lambda)(\lambda + \sigma^2)}{\lambda} \simeq (\lambda + \sigma^2)$

It is possible to estimate consistently the $(\lambda_{k,N})_{k=1,\ldots,i}$ from the $(\hat{\lambda}_{k,N})_{k=1,\ldots,i}$

For $k = 1, \ldots, i$, it holds that

$$\lambda_{k,N} - g_N(\hat{\lambda}_{k,N}) \rightarrow 0$$

where g_N is the inverse of function $\lambda \rightarrow \frac{(\sigma^2 c_N + \lambda)(\lambda + \sigma^2)}{\lambda}$

Main result on the eigenvectors

Theorem 2: Benaych-Georges and Nadakuditi, 2011

- Assume the setting of Theorem 1. Assume in addition that $\rho_1 > \rho_2 > \cdots > \rho_i \; (> \sigma^2 \sqrt{c_*})$.
- Then for $k \leq i$, for any sequences $\mathbf{b}_{1,N}, \mathbf{b}_{2,N}$ of deterministic $M \times 1$ vectors such that $\sup_N \|\mathbf{b}_{j,N}\| < \infty$, $j = 1, 2$,

$$\mathbf{b}_{1,N}^* \left(\hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* - h(\gamma_{k,N}) \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{b}_{2,N} \xrightarrow[N \to \infty]{a.s.} 0$$

  where $h(x)$ is a known function depending on σ^2 and c_N, verifying $0 < h(\gamma_{k,N}) < 1$.

Comments on Theorem II.

It is possible to estimate consistently $\mathbf{b}_N^* \left(\sum_{k=1}^{i} \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{b}_N$

- $|\mathbf{b}_N^* \mathbf{u}_{k,N}|^2 - \frac{1}{h(\gamma_{k,N})} |\mathbf{b}_N^* \hat{\mathbf{u}}_{k,N}|^2 \rightarrow 0$
- As $\hat{\lambda}_{k,N} - \gamma_{k,N} \rightarrow 0$, we have $\mathbf{b}_N^* \left(\sum_{k=1}^{i} \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{b}_N - \mathbf{b}_N^* \left(\sum_{k=1}^{i} \frac{1}{h(\hat{\lambda}_{k,N})} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{b}_N^* \rightarrow 0$

If $\rho_K > \sigma^2 \sqrt{c_*}$, or equivalently if $i = K$, it is possible to estimate consistently quadratic forms of the projection matrix on the "signal subspace"

$\rho_K > \sigma^2 \sqrt{c_*}$ refered to as the "Signal Subspace Separation Condition".

Nearly equivalent to $\lambda_{K,N} > \sigma^2 \sqrt{c_N}$ if N is large enough.

Comments on Theorem II.

Theorem II implies that if $k \neq l \leq i$, then $\mathbf{u}_{l,N}^* \hat{\mathbf{u}}_{k,N} \rightarrow 0$

- Consider $\mathbf{b}_{1,N} = \mathbf{b}_{2,N} = \mathbf{u}_{l,N}$

Theorem II does not imply that if $k \leq i$, $\hat{\mathbf{u}}_{k,N}$ is a good estimate of $\mathbf{u}_{k,N}$.

- $\mathbf{b}_{1,N} = \mathbf{b}_{2,N} = \mathbf{u}_{k,N}$ yields to $\mathbf{u}_{k,N}^* \hat{\mathbf{u}}_{k,N} - \sqrt{h(\gamma_{k,N})} \rightarrow 0$ (up to a modulus 1 coefficient)
- $0 < h(\gamma_{k,N}) < 1$ can be written as

$$h(\gamma_{k,N}) = \frac{1 - \left(\sigma^2 \sqrt{c_N} / \lambda_{k,N}\right)^2}{1 + \sigma^2 c_N / \lambda_{k,N}}$$

- If $c_N \simeq 0$, $h(\gamma_{k,N}) \simeq 1$ and $\mathbf{u}_{k,N}^* \hat{\mathbf{u}}_{k,N} \simeq 1$
- If $\lambda_{k,N}$ is close from $\sigma^2 \sqrt{c_N}$, $\mathbf{u}_{k,N}^* \hat{\mathbf{u}}_{k,N} \simeq 0$

Comments on Theorem II.

Assume $\rho_K > \sigma^2\sqrt{c_*}$, nearly equivalent to $\lambda_{K,N} > \sigma^2\sqrt{c_N}$.

- $\hat{\mathbf{U}}_N = (\hat{\mathbf{U}}_{1,N}, \hat{\mathbf{U}}_{2,N})$, with $\hat{\mathbf{U}}_{1,N}$ $M \times K$ eigenvectors associated to the K greatest eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$.
- Then, it holds that

$$\mathbf{U}_N^* \hat{\mathbf{U}}_{1,N} \simeq \mathrm{Diag}\left(\sqrt{h(\gamma_{1,N})}, \ldots, \sqrt{h(\gamma_{K,N})}\right)$$

up to a diagonal $K \times K$ matrix with unit norm entries

Some insights on the proof: eigenvalues I

$\lambda > \sigma^2(1+\sqrt{c_*})^2$ eigenvalue of $\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N}$ iff $\det\left(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - \lambda\,\mathbf{I}\right) = 0$

SVD of the signal matrix: $\frac{\mathbf{A}_N\mathbf{S}_N}{\sqrt{N}} = \mathbf{U}_N\mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*$

$$\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - \lambda\mathbf{I} = \frac{\mathbf{V}_N\mathbf{V}_N^*}{N} - \lambda\,\mathbf{I}$$

$$+ (\mathbf{U}_N, \frac{\mathbf{V}_N}{\sqrt{N}}\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2}) \begin{pmatrix} \mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\frac{\mathbf{V}_N^*}{\sqrt{N}} \end{pmatrix}$$

As $\lambda > \sigma^2(1+\sqrt{c_*})^2$, $\mathbf{Q}_N(\lambda) = \left(\frac{\mathbf{V}_N\mathbf{V}_N^*}{N} - \lambda\,\mathbf{I}_M\right)^{-1}$ and $\tilde{\mathbf{Q}}_N(\lambda) = \left(\frac{\mathbf{V}_N^*\mathbf{V}_N}{N} - \lambda\,\mathbf{I}_N\right)^{-1}$ are well defined.

Some insights on the proof: eigenvalues II

$\det\left(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - \lambda\,\mathbf{I}_M\right) = 0$ iff

$$\det\left(\mathbf{I}_M + \mathbf{Q}_N(\lambda)(\mathbf{U}_N, \frac{\mathbf{V}_N}{\sqrt{N}}\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2})\begin{pmatrix}\mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0\end{pmatrix}\begin{pmatrix}\mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\frac{\mathbf{V}_N^*}{\sqrt{N}}\end{pmatrix}\right) = 0 \tag{1}$$

or equivalently, iff

$$\det\left[\mathbf{I}_{2K} + \begin{pmatrix}\mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\frac{\mathbf{V}_N^*}{\sqrt{N}}\end{pmatrix}\mathbf{Q}_N(\lambda)(\mathbf{U}_N, \frac{\mathbf{V}_N}{\sqrt{N}}\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2})\begin{pmatrix}\mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0\end{pmatrix}\right] = 0$$

Some insights on the proof: eigenvalues III

$t_N(z)$ Stieltjes trans. of $\mathrm{MP}(\sigma^2, c_N)$, $\tilde{t}_N(z) = c_N t_N(z) - (1 - c_N)/z$

Use results concerning the behaviour of bilinear forms of $\mathbf{Q}_N(\lambda)$ and $\tilde{\mathbf{Q}}_N(\lambda)$

- $\mathbf{U}_N^* \mathbf{Q}_N(\lambda) \mathbf{U}_N \simeq t_N(\lambda)\, \mathbf{I}_K$
- $\tilde{\mathbf{U}}_N^* \frac{\mathbf{V}_N^*}{\sqrt{N}} \mathbf{Q}_N(\lambda) \mathbf{U}_N \simeq 0$
- $\tilde{\mathbf{U}}_N^* \frac{\mathbf{V}_N^*}{\sqrt{N}} \mathbf{Q}_N(\lambda) \frac{\mathbf{V}_N}{\sqrt{N}} \tilde{\mathbf{U}}_N = \tilde{\mathbf{U}}_N^* (\mathbf{I}_N + z \tilde{\mathbf{Q}}_N(z)) \tilde{\mathbf{U}}_N \simeq (1 + \lambda \tilde{t}_N(\lambda))\, \mathbf{I}_K$

Limit form of equation $\det \left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} - \lambda\, \mathbf{I} \right) = 0$

$$\det \left[\mathbf{\Lambda}_N - w_N(\lambda)\, \mathbf{I}_K \right] \simeq 0, \; w_N(\lambda) = (\lambda t_N(\lambda) \tilde{t}_N(\lambda))^{-1}$$

Conclusion follows from the observation that $\lambda \to w_N(\lambda)$ increases from $\sigma^2 \sqrt{c_N}$ to $+\infty$ when λ increases from $\sigma^2 (1 + \sqrt{c_N})^2$ to $+\infty$

Some insights on the proof: eigenvectors

$$\hat{\mathbf{\Pi}}_{k,N} = \hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^* = \frac{1}{2i\pi}\int_{\mathcal{C}_k}\left(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - z\mathbf{I}\right)^{-1}\, dz$$

where $\mathcal{C}_k$ is a contour enclosing only $\gamma_{k,N}$ and thus eigenvalue $\hat{\lambda}_{k,N}$

- Express $(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - z\mathbf{I})^{-1}$ en terms of $\mathbf{Q}_N(z) = (\frac{\mathbf{V}_N\mathbf{V}_N^*}{N} - z\mathbf{I})^{-1}$, $\tilde{\mathbf{Q}}_N(z)$, $\frac{\mathbf{V}_N}{\sqrt{N}}$, and of $\mathbf{U}_N, \tilde{\mathbf{U}}_N, \mathbf{\Lambda}_N$
- Use the asymptotic behaviour of the bilinear forms of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$
- Prove that bilinear forms of $(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - z\mathbf{I})^{-1}$ have the same behaviour than the bilinear forms of matrix $\mathbf{T}_N(z) = \left(-z(1+\sigma^2\tilde{t}_N(z)) + \frac{\mathbf{A}_N\left(\mathbf{S}_N\mathbf{S}_N^*/N\right)\mathbf{A}_N^*}{1+\sigma^2 c_N t_N(z)}\right)^{-1}$
- Evaluate the integral using the residue theorem
- Conclude

Testing $K = 0$ versus $K = 1$.

Nadakuditi-Edelmann (IEEE-SP 2008), Nadler (IEEE-SP 2010), Bianchi-Debbah-Maeda-Najim (IEEE-IT 2011) when $(s_n)_{n=1,\ldots,N}$ is an i.i.d. complex Gaussian sequence.

Hypothesis test:
$$\begin{cases} \textbf{H0} & : \quad \mathbf{Y}_N = \mathbf{V}_N & \text{(Noise)} \\ \textbf{H1} & : \quad \mathbf{Y}_N = \mathbf{a}_N \, \mathbf{s}_N + \mathbf{V}_N & \text{(Info+Noise)} \end{cases}$$

$$\lambda_{1,N} = \lambda_{max}\left((\mathbf{a}_N \mathbf{s}_N \mathbf{s}_N^* \mathbf{a}_N^*)/N\right) = \|\mathbf{a}_N\|^2 \, \frac{1}{N} \sum_{n=1}^{N} |s_n|^2 \to \rho$$

$\frac{\lambda_{1,N}}{\sigma^2}$ matched filter SNR output

Testing $K = 0$ versus $K = 1$.

Generalized Likelihood Ratio Test (GLRT)

$$T_N = \frac{\hat{\lambda}_{1,N}}{\frac{1}{M}\operatorname{tr}\left(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N}\right)}$$

Analysis of T_N under each hypothesis.

- Asymptotic analysis of T_N provides interesting insights.

Under either **H0** or **H1**

$$\frac{1}{M}\operatorname{tr}\left(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N}\right) \xrightarrow[N\to\infty]{\text{a.s.}} \sigma^2.$$

Testing $K = 0$ versus $K = 1$.

Under H0, $T_N \simeq (1 + \sqrt{c_N})^2$

Under H1

- If $\rho > \sigma^2 \sqrt{c_*}$ $\quad (\frac{\lambda_{1,N}}{\sigma^2} > \sqrt{c}_N)$, then

$$T_N \simeq \frac{\left(\sigma^2 c_N + \lambda_{1,N}\right)\left(\lambda_{1,N} + \sigma^2\right)}{\sigma^2 \, \lambda_N} > (1 + \sqrt{c_N})^2$$

- If $\rho < \sigma^2 \sqrt{c_*}$ $\quad (\frac{\lambda_{1,N}}{\sigma^2} < \sqrt{c}_N)$, then

$$T_N \simeq (1 + \sqrt{c_N})^2.$$

Testing $K = 0$ versus $K = 1$

Remarks

- $\rho > \sigma^2 \sqrt{c_*}$ provides the **limit of detectability** by the GLRT.
- False Alarm Probability can be evaluated with the help of the Tracy-Widom law.
- If sequence $(s_n)_{n=1,\ldots,N}$ is known (training sequence), no limit of detectability

Testing $K = 0$ versus $K = K_0$.

Generalized Likelihood Ratio Test (GLRT)

$$T_N = \frac{\sum_{k=1}^{K_0} \hat{\lambda}_{k,N}}{\frac{1}{M} \operatorname{tr}\left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}\right)}$$

Under either **H0** or **H1**

$$\frac{1}{M} \operatorname{tr}\left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}\right) \xrightarrow[N \to \infty]{\text{a.s.}} \sigma^2.$$

Testing $K = 0$ versus $K = K_0$.

Under H0, $T_N \simeq K_0 \, (1 + \sqrt{c_N})^2$

Under H1

- If $\rho_1 \leq \sigma^2 \sqrt{c_*}$ then

$$T_N \simeq K_0 \, (1 + \sqrt{c_N})^2.$$

- If $\rho_k > \sigma^2 \sqrt{c_*} \quad k = 1, \ldots, i$ and $\rho_{i+1} \leq \sigma^2 \sqrt{c_*}$ then

$$T_N \simeq \frac{\sum_{k=1}^{i} \gamma_{k,N} + (K_0 - i)\sigma^2 (1 + \sqrt{c_N})^2}{\sigma^2} > K_0 \, (1 + \sqrt{c_N})^2$$

Dimension reduction via principal component analysis I.

Project the M-dimensional observations $\mathbf{y}_n$ on the eigenspace associated to the K largest eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$

- $\hat{\mathbf{U}}_N = (\hat{\mathbf{U}}_{1,N}, \hat{\mathbf{U}}_{2,N})$, with $\hat{\mathbf{U}}_{1,N}$ $M \times K$ and $\hat{\mathbf{U}}_{2,N}$ $M \times (M-K)$ matrices
- $\mathbf{z}_n = \hat{\mathbf{U}}_{1,N}^* \mathbf{y}_n$ is the K–dimensional reduced size observation

Assume for simplicity that $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \simeq \mathbf{I}_K$, so that $\mathbf{A}_N = \mathbf{U}_N \mathbf{\Lambda}_N^{1/2} \mathbf{\Theta}_N^*$

Analysis of the possible SNR loss: $c_N \simeq 0$

- $c_N \simeq 0$ implies that $\hat{\mathbf{U}}_{1,N} \simeq \mathbf{U}_N$ and $\hat{\mathbf{U}}_{1,N}^* \mathbf{U}_N \simeq \mathbf{I}_K$
- $\mathbf{z}_n \simeq \mathbf{\Lambda}_N^{1/2} \mathbf{\Theta}_N^* \mathbf{s}_n + \mathbf{w}_n$
- $\mathbf{w}_n = \hat{\mathbf{U}}_{1,N}^* \mathbf{v}_n$ has covariance matrix $\sigma^2 \mathbf{I}_K$

No SNR loss if $c_N \simeq 0$

Dimension reduction via principal component analysis II.

Assume $\rho_K > \sigma^2\sqrt{c_*}$ (nearly equivalent to $\lambda_{K,N} > \sigma^2\sqrt{c_N}$).

- $\hat{\lambda}_{k,N} \simeq \gamma_{k,N} = \frac{\left(\sigma^2 c_N + \lambda_{k,N}\right)\left(\lambda_{k,N}+\sigma^2\right)}{\lambda_{k,N}}$ for $k = 1, \ldots, K$.

Analysis of the possible SNR loss: $c_N \rightarrow c_*$

- $\hat{\mathbf{U}}_{1,N}^* \mathbf{U}_N \simeq \mathrm{Diag}(\sqrt{h(\gamma_{1,N})}, \ldots, \sqrt{h(\gamma_{K,N})})$
- $h(\gamma_{k,N}) = \frac{1-\left(\sigma^2\sqrt{c_N}/\lambda_{k,N}\right)^2}{1+\sigma^2 c_N/\lambda_{k,N}} < 1$
- $\mathbf{z}_n = \mathrm{Diag}\left(\sqrt{\lambda_{1,N}}\sqrt{h(\gamma_{1,N})}, \ldots, \sqrt{\lambda_{K,N}}\sqrt{h(\gamma_{K,N})}\right) \mathbf{\Theta}_N^* \mathbf{s}_n + \mathbf{w}_n$

SNR loss on each eigenvalue: the price to be paid in the context $\frac{M}{N}$ non negligible

Source localization using the subspace method.

- Mestre-Lagunas (IEEE-SP 2008) when the source signals are i.i.d. gaussian independent sequences (use of the zero-mean correlated model).
- In the context of Information plus Noise models, see Vallet-Loubaton-Mestre (IEEE-IT 2012), Hachem-Loubaton-Mestre-Najim-Vallet (J. Multivariate Analysis 2013), Vallet-Mestre-Loubaton (IEEESP 2015)

Subspace estimation

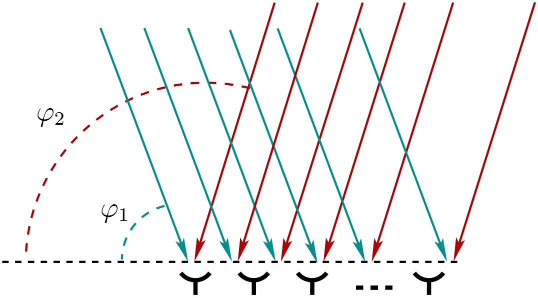
- $\mathbf{\Pi}_N = \sum_{k=1}^{K} \mathbf{\Pi}_{k,N}$ orthogonal projection on the column space of $\mathbf{A}$, $\mathbf{\Pi}_N^{\perp}$ the orthogonal projection on $[\mathrm{sp}(\mathbf{A})]^{\perp}$
- Consistent estimation of $\mathbf{b}_N^* \mathbf{\Pi}_N^{\perp} \mathbf{b}_N$, or equivalently of $\mathbf{b}_N^* \mathbf{\Pi}_N \mathbf{b}_N$, $\mathbf{b}_N$ uniformly bounded deterministic vector.

Source localization.

Problem

K radio sources send their signals to a uniform array of M antennas during N signal snapshots.

Estimate arrival angles $\varphi_1, \ldots, \varphi_K$



Example with two sources

Source localization with the subspace method (MUSIC)

Model.

- $\mathbf{Y}_N = \mathbf{A}_N \mathbf{S}_N + \mathbf{V}_N$
- $\mathbf{A}_N = \left[\mathbf{a}_N(\varphi_1) \quad \cdots \quad \mathbf{a}_N(\varphi_K)\right]$ with $\mathbf{a}_N(\varphi) = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 \\ e^{\imath\varphi} \\ \vdots \\ e^{\imath(M-1)\varphi} \end{bmatrix}$

MUSIC algorithm principle

- $\mathbf{a}_N(\varphi)^* \mathbf{\Pi}_N^{\perp} \mathbf{a}_N(\varphi) = 0 \quad \Leftrightarrow \quad \varphi \in \{\varphi_1, \ldots, \varphi_K\}$
- Estimate $\mathbf{a}_N(\varphi)^* \mathbf{\Pi}_N^{\perp} \mathbf{a}_N(\varphi)$ for each φ, and evaluate the arguments of the local minima of the estimate w.r.t. φ.
- Traditional estimate : $\mathbf{a}_N(\varphi)^* \left(\sum_{k=K+1}^{M} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^*\right) \mathbf{a}_N(\varphi)$.

Improved estimation of the cost function I

Subspace separation condition

The source number K is fixed and for all $k \in \{1, \ldots, K\}$, $\lambda_{k,N} \to \rho_k$, where $\rho_1 > \ldots > \rho_K > \sigma^2 \sqrt{c_*}$.

Traditional estimate $\mathbf{a}_N(\varphi)^* \left(\mathbf{I} - \sum_{k=1}^{K} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{a}_N(\varphi)$ converges to:

$$\mathbf{a}_N(\varphi)^* \left(\mathbf{I} - \sum_{k=1}^{K} h(\gamma_{k,N}) \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{a}_N(\varphi)$$

where $h(\gamma_{k,N}) = \frac{\lambda_{k,N}^2 - \sigma^4 c_N}{\lambda_{k,N}(\lambda_{k,N} + \sigma^2 c_N)}$ (recall that $\gamma_{k,N} = \frac{(\lambda_{k,N}+\sigma^2)(\lambda_{k,N}+\sigma^2 c)}{\lambda_{k,N}}$).

Improved estimation of the cost function I

Subspace separation condition

The source number K is fixed and for all $k \in \{1, \ldots, K\}$, $\lambda_{k,N} \to \rho_k$, where $\rho_1 > \ldots > \rho_K > \sigma^2\sqrt{c_*}$.

Traditional estimate $\mathbf{a}_N(\varphi)^* \left(\mathbf{I} - \sum_{k=1}^K \hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\right) \mathbf{a}_N(\varphi)$ converges to:

$$\underbrace{\mathbf{a}_N(\varphi)^*\mathbf{\Pi}_N^{\perp}\mathbf{a}_N(\varphi)}_{\text{MUSIC cost function}} + \underbrace{\mathbf{a}_N(\varphi)^* \left(\sum_{k=1}^{K} [1 - h(\gamma_{k,N})]\ \mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\right) \mathbf{a}_N(\varphi)}_{\text{Bias}}$$

where $h(\gamma_{k,N}) = \frac{\lambda_{k,N}^2 - \sigma^4 c_N}{\lambda_{k,N}(\lambda_{k,N}+\sigma^2 c_N)}$ (recall that $\gamma_{k,N} = \frac{(\lambda_{k,N}+\sigma^2)(\lambda_{k,N}+\sigma^2 c)}{\lambda_{k,N}}$).

Improvement

Need to apply some correction (Theorem II) to recover consistency.

Improved estimation of the cost function II

Consistent estimate of $\mathbf{a}_N(\varphi)^* \mathbf{\Pi}^\perp \mathbf{a}_N(\varphi)$

$$\mathbf{a}_N(\varphi)^* \left(\mathbf{I} - \sum_{k=1}^{K} \frac{\hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^*}{h(\hat{\lambda}_{k,N})} \right) \mathbf{a}_N(\varphi)$$

Stronger result - Uniform convergence

$$\sup_{\varphi \in (-\pi, \pi]} \left| \mathbf{a}_N(\varphi)^* \Pi_N^\perp \mathbf{a}_N(\varphi) - \left(1 - \sum_{k=1}^{K} \frac{|\mathbf{a}_N(\varphi)^* \hat{\mathbf{u}}_{k,N}|^2}{h(\hat{\lambda}_{k,N})} \right) \right| \xrightarrow[N \to \infty]{a.s.} 0$$

Remark. Uniform consistency of the cost function estimator over φ is required to study the asymptotic behaviour of the DoA estimates.

Asymptotic behaviour of the improved DoA estimates

Widely spaced DoA scenario

- **Widely spaced DoA.** $\varphi_1, \ldots, \varphi_K$ are fixed w.r.t. N. Implies that $\mathbf{u}_{k,N} \simeq \mathbf{a}_N(\phi_k)$ for $k = 1, \ldots, K$.
- **Uncorrelated sources.** $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$ converge to $\mathrm{diag}(\rho_1, \ldots, \rho_K)$.
- **SNR condition.** $\rho_K > \sigma^2 \sqrt{c_*}$

$\Rightarrow$ **The subspace separation is satisfied.**

M-Consistency - Widely spaced DoA

For all $k \in \{1, \ldots, K\}$,

$$M\left(\hat{\varphi}_{k,N} - \varphi_k\right) \xrightarrow[N\to\infty]{a.s.} 0$$

Asymptotic normality - Widely spaced DoA

For all $k \in \{1, \ldots, K\}$,

$$M^{3/2} \left(\hat{\varphi}_{k,N} - \varphi_k \right) \xrightarrow[N \to \infty]{\mathcal{D}} \mathcal{N} \left(0, \frac{6\sigma^2(\rho_k + \sigma^2)}{\rho_k^2 - \sigma^4 c_*} \right)$$

Comments

- **Both results proved in Hachem et. al '12**
- **If a source power is close to $\sigma^2 \sqrt{c_*}$, the corresponding MSE increases.**
- **Results can be extended to the case of correlated sources.**

Sketch of proof

- **Improved estimate.** $\hat{\eta}_N(\varphi) = 1 - \sum_{k=1}^{K} \frac{|\mathbf{a}_N(\varphi)^* \hat{\mathbf{u}}_{k,N}|^2}{h(\hat{\lambda}_{k,N})}$.
- **Taylor expansion.** As $M(\hat{\varphi}_k - \varphi_k) \to 0$ a.s., we have

$$M^{3/2}\left(\hat{\varphi}_{k,N} - \varphi_k\right) = -\frac{\frac{1}{\sqrt{M}}\hat{\eta}'_N(\varphi_k)}{\frac{1}{M^2}\hat{\eta}_N^{(2)}(\varphi_k)} + o_{\mathbb{P}}(1).$$

- **1st order.**

$$\frac{1}{M^2}\hat{\eta}_N^{(2)} = 2\frac{\mathbf{a}'_N(\varphi_k)^*}{M}\Pi_N^{\perp}\frac{\mathbf{a}'_N(\varphi_k)^*}{M} + o_{\mathbb{P}}(1).$$

- **2nd order.** Need to derive CLT on the bilinear form

$$\frac{1}{\sqrt{M}}\hat{\eta}'_N(\varphi_k) = 2\sqrt{M}\mathrm{Re}\left[\frac{\mathbf{a}'_N(\varphi_k)^*}{M}\left(\mathbf{I} - \sum_{k=1}^{K}\frac{\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*}{h(\hat{\lambda}_{k,N})}\right)\mathbf{a}_N(\varphi_k)\right].$$

CLT for bilinear forms.

Theorem

Let $(\mathbf{b}_{1,N})$, $(\mathbf{b}_{2,N})$ *two deterministic sequences of unit norm vectors. If*

- $c_N = c_* + o(N^{-1/2})$,
- $\liminf_N \|\Pi_N \mathbf{b}_{1,N}\| > 0$,

there exists a deterministic bounded sequence (ξ_N) *s.t.* $\liminf_N \xi_N > 0$ *and*

$$\sqrt{\frac{N}{\xi_N}} \mathrm{Re}\left(\mathbf{b}_{1,N}^* \left(\mathbf{I} - \sum_{k=1}^{K} \frac{\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*}{h(\hat{\lambda}_{k,N})} \right) \mathbf{b}_{2,N} - \mathbf{b}_{1,N}^* \Pi_N^{\perp} \mathbf{b}_{2,N} \right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,1).$$

Remark

The rate $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$ does not hold anymore if $\mathbf{b}_{1,N}, \mathbf{b}_{2,N}$ belong to the noise subspace.

Sketch of proof

- **Integral representation.** If $\mathcal{C}$ is a contour enclosing $\gamma_{1,N}, \ldots, \gamma_{K,N}$ (and thus $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$) and not 0,

$$\sum_{k=1}^{K} \frac{\mathbf{b}_{1,N}^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{b}_{2,N}}{h(\hat{\lambda}_{k,N})} = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\mathbf{b}_{1,N}^* \left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} - z\mathbf{I} \right)^{-1} \mathbf{b}_{2,N}}{h(z)} \mathrm{d}z$$

- **CLT for quadratic forms** The random process

$$z \mapsto \mathbf{b}_{1,N}^* \left(\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} - z\mathbf{I} \right)^{-1} \mathbf{b}_{2,N}$$

  defined on the compact $\mathcal{C}$ converges in distribution to a continuous Gaussian process, with fluctuations of the order $\mathcal{O}\left(\sqrt{\frac{1}{N}}\right)$.

- **Transfer to the integral.** Integral of a continuous Gaussian process = Gaussian R.V.

Behaviour of standard MUSIC

- **Cost function.** Uniformly on φ,

$$\mathbf{a}_N(\varphi)^* \, \hat{\Pi}_N^{\perp} \, \mathbf{a}_N(\varphi) \approx 1 - \sum_{k=1}^{K} h(\gamma_{k,N}) \, |\mathbf{a}_N(\varphi)^* \mathbf{a}(\varphi_k)|^2$$

- **Minimizers.** The asymptotic cost function admits $\varphi_1, \ldots, \varphi_K$ as unique minimizers.
- **DoA estimates.** We can prove that traditional MUSIC DoA estimates satisfy exactly the same 1st and 2nd order results that the improved estimates.

Remarks

- $\Rightarrow$ **No improvement in this scenario!**
- **The basic smoothed periodogram also lead to consistent estimators with same rate of convergence, but subspace separation condition not needed.**

Closely spaced DoA scenario

- $K = 2$ sources
- **Closely spaced DoA.** $\varphi_{2,N} = \varphi_{1,N} + \frac{\alpha}{M}$
- **Uncorrelated sources** $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \to \mathbf{I}_2$, which implies that

$$\lambda_{1,N} \to \rho_1 = 1 + |\mathrm{sinc}(\alpha/2)| \quad \text{and} \quad \lambda_{2,N} \to \rho_2 = 1 - |\mathrm{sinc}(\alpha/2)|$$

- **Subspace separation condition.** $1 - |\mathrm{sinc}(\alpha/2)| > \sigma^2 \sqrt{c_*}$.

M-Consistency - Closely spaced DoA

For all $k \in \{1, 2\}$,

$$M\left(\hat{\varphi}_{k,N} - \varphi_{k,N}\right) \xrightarrow[N\to\infty]{a.s.} 0$$

Asymptotic normality - Closely spaced DoA

For all $k \in \{1, 2\}$

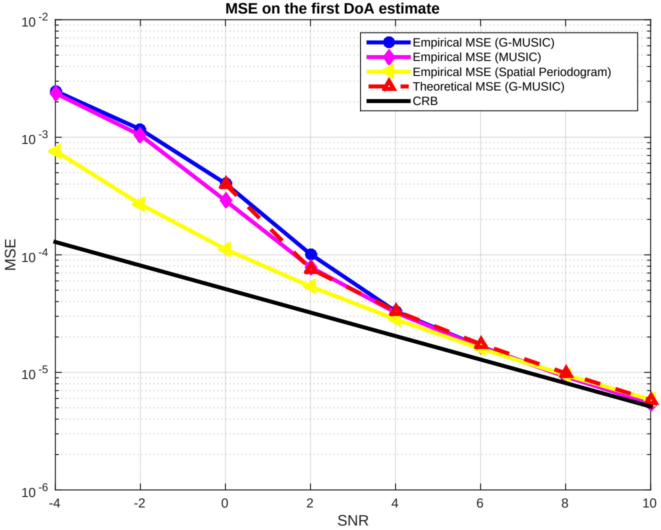
$$\frac{M^{3/2}}{\sqrt{\tilde{\xi}_{k,N}}} \left(\hat{\varphi}_{k,N} - \varphi_{k,N}\right) \xrightarrow[N \to \infty]{\mathcal{D}} \mathcal{N}(0, 1).$$

where $(\xi_{k,N})$ is a bounded deterministic sequence s.t. $\liminf_N \xi_{k,N} > 0$.

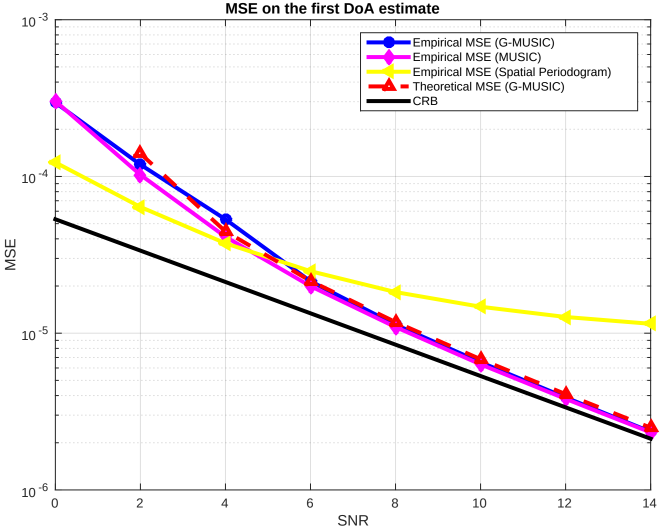
Comments

- **The improved MUSIC method is still able to asymptotically separate closely spaced DoA.**
- **Smoothed periodogram and standard MUSIC are not M-consistent anymore.**

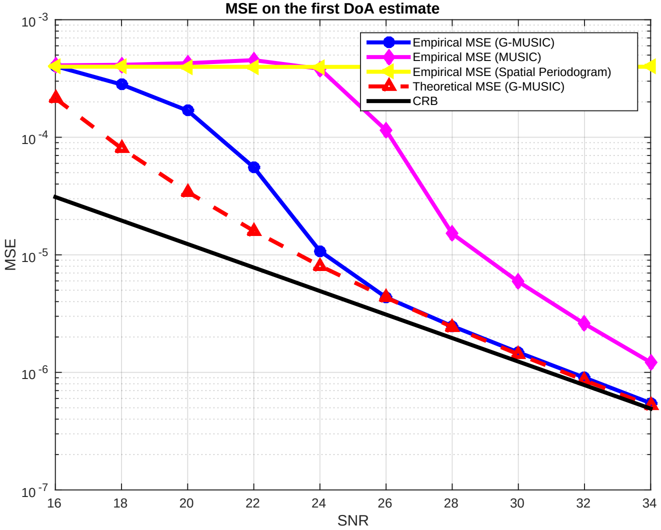
$K = 2$, $M = 40$, $N = 80$, $\varphi_2 - \varphi_1 = 5 \times \frac{2\pi}{M}$, uncorrelated sources.



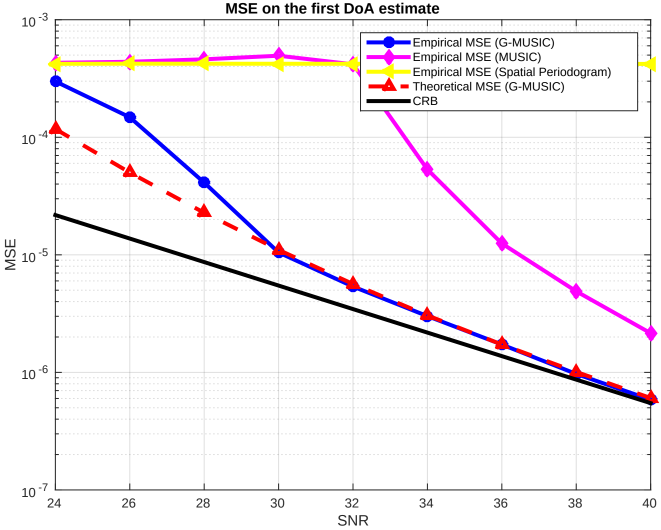
$K = 2$, $M = 40$, $N = 80$, $\varphi_2 - \varphi_1 = 5 \times \frac{2\pi}{M}$, correlated sources.



$K = 2$, $M = 40$, $N = 80$, $\varphi_2 - \varphi_1 = \frac{1}{4} \times \frac{2\pi}{M}$, uncorrelated sources.



$K = 2,\ M = 40, N = 20,\ \varphi_2 - \varphi_1 = \frac{1}{4} \times \frac{2\pi}{M}$, uncorrelated sources.

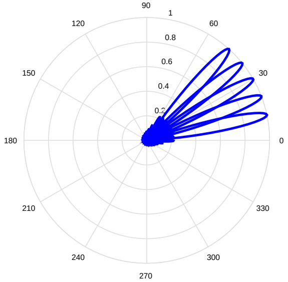


Other frequently used methods

Beamspace MUSIC

- **Idea.** Prefiltering the data to focus the array onto an angular sector Θ where the DoA are located.
- **DFT Beamformer.** Form L orthonormal beams $\mathbf{a}(\psi_{1,N}), \ldots, \mathbf{a}(\psi_{L,N})$ with

$$\{\psi_{1,N}, \ldots, \psi_{L,N}\} = \left\{-\pi + \frac{2\pi(m-1)}{M} : m = 1, \ldots, M\right\} \cap \Theta.$$



Filtered signal

$$\tilde{\mathbf{Y}}_N = \mathbf{B}_N^* \mathbf{Y}_N$$
$$= \tilde{\mathbf{A}}_N \mathbf{S}_N + \tilde{\mathbf{V}}_N,$$

where

- $\mathbf{B}_N = [\mathbf{a}(\psi_{1,N}), \ldots, \mathbf{a}(\psi_{L,N})]$
- $\tilde{\mathbf{A}}_N = [\tilde{\mathbf{a}}_N(\varphi_1), \ldots, \tilde{\mathbf{a}}_N(\varphi_K)]$, with $\tilde{\mathbf{a}}_N(\varphi_k) = \mathbf{B}_N^* \mathbf{a}_N(\varphi_k)$.
- $\tilde{\mathbf{V}}_N = \mathbf{B}_N^* \mathbf{V}_N$ has i.i.d $\mathcal{CN}(0, \sigma^2)$ entries.

Beamspace MUSIC algorithm

Estimate the DoA as the K deepest minima of

$$\varphi \mapsto \tilde{\mathbf{a}}_N(\varphi)^* \tilde{\Pi}_N^{\perp} \tilde{\mathbf{a}}_N(\varphi),$$

where $\tilde{\Pi}_N^{\perp}$ is the noise projector estimate based on $\tilde{\mathbf{Y}}_N$.

Dimensionality reduction - L scales with N

If Θ is fixed w.r.t. N,

$$\frac{L}{N} \to d_* = \frac{|\Theta|}{2\pi} c_* \leq c_* .$$

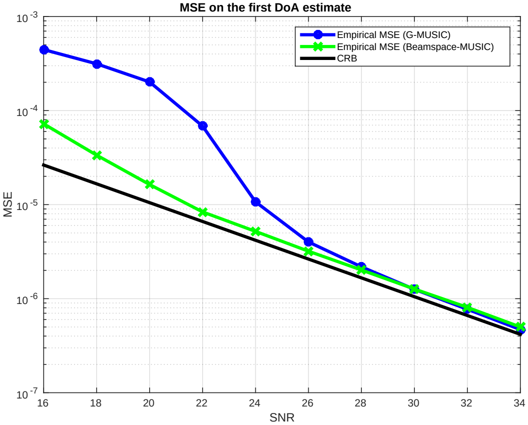
$\Rightarrow$ **The separation condition is less restrictive.**

Dimensionality reduction - L fixed w.r.t N

If L is fixed w.r.t. N (thus $|\Theta| = \mathcal{O}\left(\frac{1}{M}\right)$)

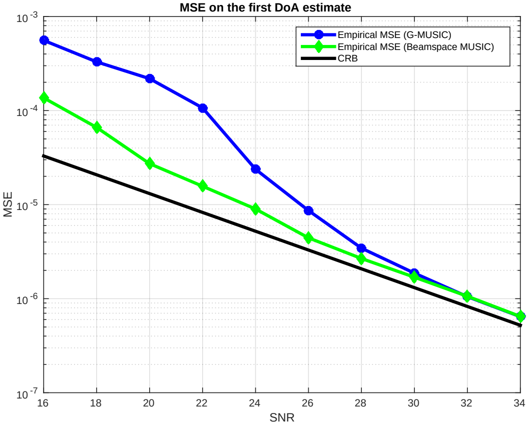
$\Rightarrow$ **The separation condition disappears and we can recover M-consistency in a closely spaced DoA scenario.**

$M = 40$, $N = 80$, $\varphi_2 - \varphi_1 = \frac{1}{4} \times \frac{2\pi}{M}$, uncorrelated sources



Focusing sector: $\Theta = [\varphi_1 - \frac{5\pi}{M}, \varphi_2 + \frac{5\pi}{M}]$

$M = 40,\ N = 80,\ \varphi_2 - \varphi_1 = \frac{1}{4} \times \frac{2\pi}{M}$, correlated sources



Focusing sector: $\Theta = [\varphi_1 - \frac{5\pi}{M}, \varphi_2 + \frac{5\pi}{M}]$

1 Problem statement

2 The narrow band array processing model
- Detailed presentation of the narrow band array processing model.
- The pure noise case: the Marcenko-Pastur distribution
- The signal plus noise case.
- Applications.

3 Wideband array processing models.
- Detailed description of the wideband array processing model
- Asymptotic behaviour of the empirical spatio-temporal covariance matrix.
- Applications.

4 Conclusion

The observed signal.

Observation: M-dimensional time series $\mathbf{y}_n$ observed from $n = 1$ to $n = N$.

- $\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n = [\mathbf{h}(z)] s_n + \mathbf{v}_n$
- $(s_n)_{n \in \mathbb{Z}}$ scalar deterministic sequence
- $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ unknown SIMO transfer function
- $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ temporally and spatially white complex Gaussian noise with variance σ^2.

Associated narrowband model with P sources.

- $\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n$
- $\mathbf{A} = (\mathbf{h}_{P-1}, \ldots, \mathbf{h}_0)$
- $\mathbf{s}_n = (s_{n-(P-1)}, s_{n-(P-1)+1}, \ldots, s_n)^T$

Does not take into account the structure of $\mathbf{s}_n$.

The extended observed signal

$(y_{k,n})_{n \in \mathbb{Z}}$ scalar signal received on sensor k.

For L well chosen, define for each n L-dimensional vector $\mathbf{y}_{k,n}^{(L)}$ by:

$\mathbf{y}_{k,n}^{(L)} = \left(y_{k,n}, y_{k,n+1}, \ldots, y_{k,n+L-1}\right)^T$ and ML–dimensional vector $\mathbf{y}_n^{(L)}$ by:

$$\mathbf{y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n}^{(L)} \\ \vdots \\ \mathbf{y}_{M,n}^{(L)} \end{pmatrix}$$

Define $ML \times N$ matrix $\mathbf{Y}_N^{(L)}$ by:

$$\mathbf{Y}^{(L)} = \left(\mathbf{y}_1^{(L)}, \ldots, \mathbf{y}_N^{(L)}\right)$$

$\mathbf{Y}^{(L)}$ is a block-Hankel matrix.

For each k, define $L \times N$ Hankel matrix $\mathbf{Y}^{(L)}_{k,N}$ by

$$\mathbf{Y}^{(L)}_{k,N} = \begin{pmatrix} y_{k,1} & y_{k,2} & \cdots & y_{k,N} \\ y_{k,2} & y_{k,3} & \cdots & y_{k,N+1} \\ y_{k,3} & \cdots & \cdots & y_{k,N+2} \\ \vdots & \vdots & \vdots & \vdots \\ y_{k,L} & y_{k,L+1} & \cdots & y_{k,N+L-1} \end{pmatrix}$$

$\mathbf{Y}^{(L)}_N$ is given by:

- $$\mathbf{Y}^{(L)}_N = \begin{bmatrix} \mathbf{Y}^{(L)}_{1,N} \\ \vdots \\ \mathbf{Y}^{(L)}_{M,N} \end{bmatrix}$$

Expression of $\mathbf{Y}_N^{(L)}$.

For each k:

- $\mathbf{Y}_{k,N}^{(L)} = \mathbf{H}_k^{(L)} \mathbf{S}_N^{(L)} + \mathbf{V}_{k,N}^{(L)}$
- where $\mathbf{H}_k^{(L)}$ is a $L \times (P+L-1)$ Toeplitz matrix and $\mathbf{S}_N^{(L)}$ is a $(P+L-1) \times N$ Hankel matrix

- $$\mathbf{Y}_N^{(L)} = \begin{pmatrix} \mathbf{H}_1^{(L)} \\ \vdots \\ \mathbf{H}_M^{(L)} \end{pmatrix} \mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)} = \mathbf{H}^{(L)} \, \mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)}$$
- $\mathbf{Y}_N^{(L)}$ can be interpreted as a low rank block-Hankel Information plus Noise random matrix model

What can be said on eigenvalues / eigenvectors of the empirical spatio-temporal covariance matrix $\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N}$?

Asymptotic behaviour of the eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$.

Asymptotic regime

- $M \to +\infty$, $N \to +\infty$, $d_N = \frac{ML}{N} \to d_*$
- L may converge towards $+\infty$ but in such a way that $\frac{L}{N} \to 0$

Theorem (PL, J. of Theo. Prob. in press)

- The empirical eigenvalue distribution of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$ has almost surely the same asymptotic behaviour than $\mathrm{MP}(\sigma^2, d_N)$
- If moreover $L = \mathcal{O}(N^\alpha)$ with $\alpha < 2/3$, nearly equivalent to $\frac{L}{M^2} \to 0$, then:
  - all the non zero eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$ lie in a neighbourhood of $[\sigma^2(1-\sqrt{d_*})^2, \sigma^2(1+\sqrt{d_*})^2]$.
  - Moreover, if $\lambda \in \mathbb{C} - [\sigma^2(1-\sqrt{d_*})^2, \sigma^2(1+\sqrt{d_*})^2]$, the bilinear forms of matrices $\mathbf{Q}_N(\lambda) = (\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N} - \lambda)^{-1}$ and $\tilde{\mathbf{Q}}_N(\lambda)$ behave as if the entries of $\mathbf{V}_N^{(L)}$ were i.i.d.

Asymptotic behaviour of the largest eigenvalues and associated eigenvectors of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N}$

Asymptotic regime

- $M \to +\infty$, $N \to +\infty$, $d_N = \frac{ML}{N} \to d_*$
- L and P do not scale with M and N

The rank of signal matrix $\mathbf{H}_N^{(L)}\mathbf{S}_N^{(L)}$ does not scale with M and N

- All the results presented above in the context of the standard low rank Information plus Noise models are still valid, but $c_N = \frac{M}{N}$ and K have to be replaced by $d_N = \frac{ML}{N}$ and $P + L - 1$

Application to the detection of signal $[\mathbf{h}(z)]s_n$ from the observations $(\mathbf{y}_n)_{n=1,\ldots,N}$.

Test based on the largest eigenvalues $(\hat{\lambda}_{k,N}^{(L)})$ of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N}$.

- $T_N^{(L)} = \frac{\sum_{k=1}^{Q+L-1} \hat{\lambda}_{k,N}^{(L)}}{\mathrm{Tr}(\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*})/N}$
- Possible to evaluate the first order behaviour of $T_N^{(L)}$ and to get insights on the effects of the choice of Q and L
- See G.T. Pham, PL, Eusipco 2015 for more details
- Consistency of the test if the largest eigenvalue $\lambda_{1,N}^{(L)}$ of $\mathbf{H}_N^{(L)} \frac{\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)}}{N} \mathbf{H}_N^{(L)*}$ is greater than $\sigma^2\sqrt{ML/N}$.
- If L increases, the detectability threshold increases, but $\lambda_{1,N}^{(L)}$ increases as well until saturation.
- The optimal choice depends on the properties of $\mathbf{h}(z)$.

Application to the detection of signal $[\mathbf{h}(z)]s_n$ from the observations $(\mathbf{y}_n)_{n=1,\ldots,N}$.

Example: Vectors $(\mathbf{h}_p)_{p=0,\ldots,P-1}$ are realizations of zero mean uncorrelated random vectors and $\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)*}/N \simeq \mathbf{I}_{P+L-1}$.

- Consistency of the test if the largest eigenvalue $\lambda_{1,N}^{(L)}$ of $\mathbf{H}_N^{(L)}\frac{\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)}}{N}\mathbf{H}_N^{(L)*}$ is greater than $\sigma^2\sqrt{ML/N}$.
- If L increases, the detectability threshold increases, but $\lambda_{1,N}^{(L)}$ increases as well until saturation.
- As M is large $\mathbf{h}_p^*\mathbf{h}_q \simeq \mu_p\,\delta_{p-q}$, $\mu_p = \mathbb{E}(\|\mathbf{h}_p\|^2)$
- If $L = 1$, $\lambda_{1,N}^{(L)} \simeq \max_{p=0}^{P-1}\mu_p$
- If $L \geq P$, $\lambda_{1,N}^{(L)} \simeq \sum_{p=0}^{P-1}\mu_p$
- If $\mu_p = \mu$ for each p,
  - for $L \leq P$, the consistency condition is $\mu \geq \frac{\sigma^2}{\sqrt{L}}\sqrt{M/N}$
  - for $L \geq P$, it is $\mu \geq \sigma^2\,(\sqrt{L}/P)\,\sqrt{M/N}$

Application to the loading factor estimation of trained spatio-temporal Wiener filters.

Observation: M-dimensional time series $\mathbf{y}_n$ observed from $n = 1$ to $n = N$.

- $\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n = [\mathbf{h}(z)]s_n + \mathbf{v}_n$
- $(s_n)_{n \in \mathbb{Z}}$ scalar deterministic sequence
- $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ unknown SIMO transfer function
- $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ temporally and spatially white complex Gaussian noise with variance σ^2.

Context.

- Training sequence $(s_n)_{n=1,\ldots,N}$ available at the receiver side, $(\mathbf{y}_n)_{n=1,\ldots,N}$ the corresponding received signal.
- Estimate ML–dimensional vector $\mathbf{g}^{(L)}$ for which $\mathbb{E}|s_n - \mathbf{g}^{(L)*}\mathbf{y}_n^{(L)}|^2$ is minimum

Application to the loading factor estimation of trained spatio-temporal Wiener filters.

Context.

- Training sequence $(s_n)_{n=1,\ldots,N}$ available at the receiver side, $(\mathbf{y}_n)_{n=1,\ldots,N}$ the corresponding received signal.
- Estimate ML–dimensional vector $\mathbf{g}^{(L)}$ for which $\mathbb{E}|s_n - \mathbf{g}^{(L)*}\mathbf{y}_n^{(L)}|^2$ is minimum
- Regularized least-squares estimate: $\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I}\right)^{-1} \left(\frac{1}{N}\sum_{n=1}^N \mathbf{y}_n^{(L)} s_n^*\right)$
- Regularization necessary if $ML > N$ and known to improve performance when ML/N is not small enough
- How to choose λ when M and N are large and of the same order of magnitude ?

Application to the loading factor estimation of trained spatio-temporal Wiener filters.

Context.

- Regularized least-squares estimate:
$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(L)} s_n^* \right)$$
- How to choose λ when M and N are large and of the same order of magnitude ?
- Questions inspired by Mestre-Lagunas IEEE SP 2006, devoted to the case $\mathbf{h}(z) = \mathbf{h}_0$ a priori known (no training sequence), temporally white but spatially correlated noise + interference with unknown covariance matrix, $L = 1$.

Maximization of the SINR provided by filter $\hat{\mathbf{g}}_\lambda^{(L)}$.

$$\text{Assume } \frac{\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)*}}{N} = \mathbf{I}_{P+L-1}$$

The SINR provided by $\hat{\mathbf{g}}_\lambda^{(L)}$ is easily seen to be

$$\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{H}_-^{(L)}\mathbf{H}_-^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2}$$

$\mathbf{h}_P^{(L)}$ column P of matrix $\mathbf{H}^{(L)}$, $\mathbf{H}_-^{(L)}$ matrix obtained from $\mathbf{H}^{(L)}$ by deleting column P.

SINR($\hat{\mathbf{g}}_\lambda^{(L)}$) is a random variable because $\hat{\mathbf{g}}_\lambda^{(L)}$ depends on the noise corrupting the signal $(\mathbf{y}_n)_{n=1,,\ldots,N}$ received during the transmission of the training sequence.

Main results

$$\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{H}_-^{(L)}\mathbf{H}_-^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2}$$

Main results: When M and N converge towards $+\infty$ at the same rate, and that P and L are fixed

- $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ converges a.s. towards a deterministic term $\phi_L(\lambda)$ depending on λ and on $\sigma^2, \mathbf{H}^{(L)}$.
- While $\mathbf{H}^{(L)}$ is unknown at the receiver side, it is possible to estimate consistently $\phi_L(\lambda)$ for each $\lambda \geq 0$ from $(\mathbf{y}_n)_{n=1,\ldots,N}$.
- λ is estimated as the argmax of the consistent estimate of $\lambda \to \phi_L(\lambda)$.

Some insights on the deterministic behaviour of the SINR.

$$\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{H}_-^{(L)}\mathbf{H}_-^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2}$$

Convergence of $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ towards a deterministic term $\phi_L(\lambda)$.

Evaluate the behaviour of

- $\mathbf{u}^* \, \hat{\mathbf{g}}_\lambda^{(L)}$ for each deterministic ML–dimensional vector $\mathbf{u}$.
- $\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2$

Expression of $\hat{\mathbf{g}}_\lambda^{(L)}$.

$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(L)} s_n^* \right)$$

Matrix $\left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1}$ coincides with the resolvent of matrix $\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N}$ at point $-\lambda$.

If $\mathbf{a}_N = \left(\frac{1}{\sqrt{N}} (s_1, s_2, \ldots, s_N) \right)^*$

$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1} \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}} \mathbf{a}_N$$

$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1} \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}} \mathbf{a}_N$$

- Possible to show that bilinear forms of $\left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1}$ have the same behaviour than if noise matrix $\mathbf{V}_N^{(L)}$ were i.i.d.
- Not sufficient: presence of $\frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}} \mathbf{a}_N$ and evaluation the behaviour of $\|\hat{\mathbf{g}}_\lambda^{(L)}\|$.
- See G.T. Pham, PL, SSP 2016 for more details.

Discussion

Assume $d_N = ML/N < 1$ and $\lambda = 0$. Denote by γ the SINR provided by the true Wiener filter:

$$\gamma = \frac{\mathbf{h}_P^{(L)*} \left(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{h}_P^{(L)}}{1 - \mathbf{h}_P^{(L)*} \left(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{h}_P^{(L)}}$$

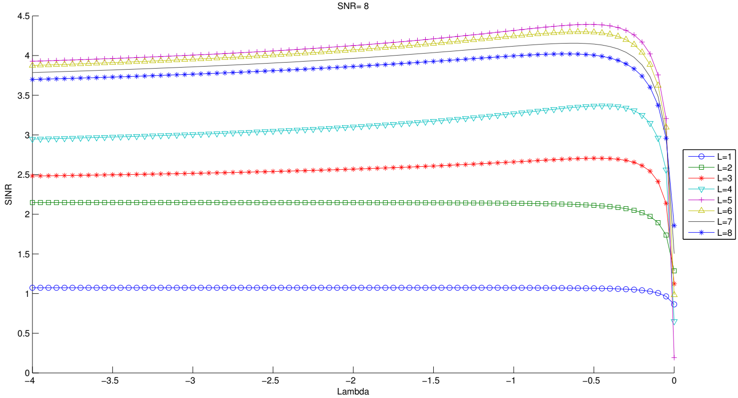
Then, the limit SINR $\phi_L(0)$ provided by $\hat{\mathbf{g}}_0^{(L)*}$ is given by

$$\phi_L(0) = \gamma \, \frac{(1 - d_N)\gamma}{\gamma + d_N}$$

SINR loss equal to $(1 - d_N) \frac{\gamma}{\gamma + d_N}$

Illustration

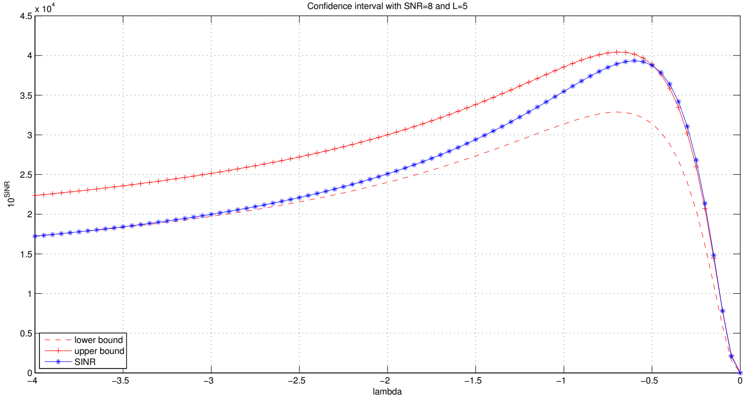
$M = 40,\ N = 200,\ P = 5,\ (\mathbf{h}_p)_{p=0,\ldots,4}$ random directional vectors



Asymptotic SINR vs λ for various values of L.

Illustration

$M = 40,\ N = 200,\ P = 5,\ L = 5,\ (\mathbf{h}_p)_{p=0,\ldots,4}$ random directional vectors



Comparison between $\phi_5(\lambda)$ and 95 per cent confidence intervals on $\mathrm{SINR}(\hat{\mathbf{g}}_\lambda^{(5)})$

Application to the analysis of subspace DoA estimation using spatial smooting schemes I.

Spatial smoothing originally designed for DoA estimation of fully correlated signals.

Also allows to use subspace method when $N << M$.

$L < M$: artificially create NL snapshots of dimension $M - L + 1$.

$$\mathcal{Y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n} & \mathbf{y}_{2,n} & \cdots & \cdots & \mathbf{y}_{L,n} \\ \mathbf{y}_{2,n} & \mathbf{y}_{3,n} & \cdots & \cdots & \mathbf{y}_{L+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{M-L+1,n} & \mathbf{y}_{M-L+2,n} & \cdots & \cdots & \mathbf{y}_{M,n} \end{pmatrix}$$

Define $(M - L + 1) \times NL$ matrix $\mathbf{Y}_N^{(L)}$ by

$$\mathbf{Y}_N^{(L)} = \left(\mathcal{Y}_1^{(L)}, \ldots, \mathcal{Y}_N^{(L)} \right)$$

Application to the analysis of subspace DoA estimation using spatial smooting schemes II.

Properties of $\mathbf{Y}_N^{(L)}$.

- $\mathbf{Y}_N^{(L)} = \mathbf{X}_N^{(L)} + \mathbf{V}_N^{(L)}$
- $\mathbf{X}_N^{(L)}$ is a rank K deterministic $(M - L + 1) \times NL$ matrix
- $\mathrm{Range}(\mathbf{X}_N^{(L)}) = \mathrm{sp}\{\mathbf{a}_{M-L+1}(\varphi_k), k = 1, \ldots, K\}$

Narrow band array processing model with $M - L + 1$ sensors and NL (correlated) observations.

Quantify the performance of subspace and improved subspace method in the high-dimensional context.

- Characterization of the K largest eigenvalues / eigenvectors of $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$

The asymptotic regime.

- $M \to +\infty$, $N = \mathcal{O}(M^{\beta})$, $1/3 < \beta \leq 1$
- $e_N = \frac{M-L+1}{NL} \to c_*$
- Implies that $L = \mathcal{O}(M^{\alpha}), 0 \leq \alpha < 2/3$, and that $\frac{M}{NL} \to e_*$

Properties of the eigenvalues of $\mathbf{V}_N^{(L)} \mathbf{V}_N^{(L)*} / NL$.

- Eigenvalue distribution has the same asymptotic behaviour than $\mathrm{MP}(\sigma^2, e_N)$
- All the eigenvalues lie in a neighbourhood of $[\sigma^2(1-\sqrt{e_*})^2, \sigma^2(1+\sqrt{e_*})^2]$
- Moreover, if $\lambda \in \mathbb{C} - [\sigma^2(1-\sqrt{e_*})^2, \sigma^2(1+\sqrt{e_*})^2]$, the bilinear forms of matrices $\mathbf{Q}_N(\lambda) = (\frac{\mathbf{V}_N^{(L)} \mathbf{V}_N^{(L)*}}{N} - \lambda \mathbf{I})^{-1}$ and $\tilde{\mathbf{Q}}_N(\lambda)$ behave as if the entries of $\mathbf{V}_N^{(L)}$ were i.i.d.

The results concerning high dimensional subsapce methods can be extended.

Subspace separation condition and largest eigenvectors of $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$.

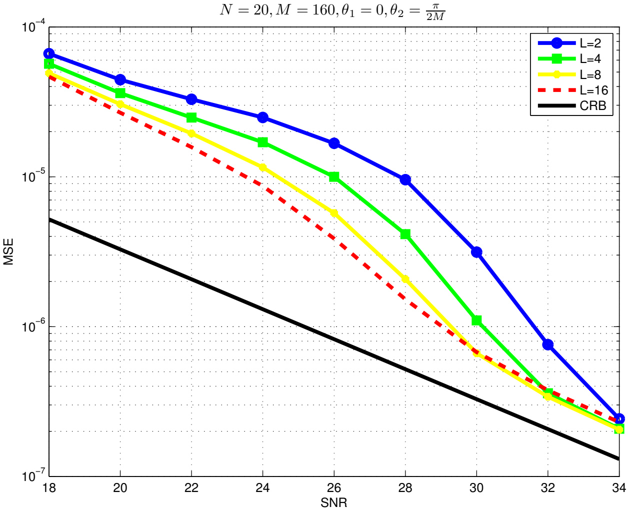
Comparison smoothed / unsmoothed when L does not converge $+\infty$.

- $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \to \mathbf{D}$, $\mathbf{D}$ diagonal
- unsmoothed: $\lambda_K\left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \sigma^2\sqrt{M/N}$
- smoothed: $\lambda_K\left(\mathbf{A}_{M-L}^*\mathbf{A}_{M-L}\mathbf{D}\right) > \frac{\sigma^2}{\sqrt{L}}\sqrt{M/N} = \sigma^2\sqrt{\frac{M}{NL}}$
- Same condition as in a standard narrow band array processing model with $M-L+1$ antennas and NL (independent) snapshots.

Discussion

- If $L << M$, $\lambda_K\left(\mathbf{A}_{M-L}^*\mathbf{A}_{M-L}\mathbf{D}\right) \simeq \lambda_K\left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right)$
- Clear improvement of the subspace separation condition if $L << M$
- If L increases too much, the diminution of the number of antennas due to the spatial smoothing becomes dominant.

Illustration



MMSE of the improved subspace estimate of θ_1 for $L = 2, 4, 8, 16$ w.r.t. SNR.

Conclusion.

- Certain classical problems have to be revisited when M and N are of the same order of magnitude
- The theoretical results that are obtained are in general reliable, even if $\frac{M}{N}$ is small
- Although rather technical, the above asymptotic technics should be widely disseminated in the community

Other related high dimensional signal processing problems.

- Consistent estimation of large covariance matrices when a priori informations are available (e.g. sparsity).
- Sparse principal component analysis.
- Different mathematical tools.

Bibliography

- N. Asendorf, R. R. Nadakuditi, "The performance of matched subspace detector that uses subspaces estimated from finite noisy training data", *IEEE Trans. SP*, vol. 61, no. 8, pp. 1972-1985, 2013.
- Z. Bai, J.W. Silverstein, "Spectral analysis of large dimensional random matrices", Springer Series in Statistics, 2nd ed., 2010.
- F. Benaych-Georges and R. R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices", *Adv. in Math.*, vol. 227, no. 1, 2011, pp. 494-521.
- F. Benaych-Georges and R. R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices", *JMVA*, 111, 2012, pp. 120-135 ArXiv e-prints, 2011.
- P. Bianchi, M. Debbah, M. Maida, J. Najim, "Performance of Statistical Tests for Source Detection using Random Matrix Theory", *IEEE Trans. IT*, 57 (4), 2011.
- R. Couillet, M. Debbah, "Random Matrix Methods for Wireless Communications", Cambridge University Press, September 2011.

- W. Hachem, P. Loubaton, J. Najim, X. Mestre, P. Vallet, "A subspace estimator for fixed rank perturbations of large random matrices", *JMVA*, 114 (2013), pp. 427-447.
- P. Loubaton, "On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices", to appear in J. of Theoretical Probability, published on line at http://link.springer.com/article/10.1007/s10959-015-0614-z, also available on Arxiv (http://arxiv.org/abs/1405.2006)
- P. Loubaton, P. Vallet, "Almost sure localization of the eigenvalues in a gaussian information plus noise model. Applications to the spiked models", *Electronic J. on Probability*, October 2011, pp. 1934-1959.
- X. Mestre, M. A. Lagunas. "Finite sample size effect on Minimum Variance beamformers: optimum diagonal loading factor for large arrays". IEEE Transactions on Signal Processing, vol. 54, no. 1, pp. 69-82, January 2006.
- X. Mestre, "Improved Estimation of Eigenvalues and Eigenvectors of Covariance Matrices Using Their Sample Estimates", *IEEE Trans. IT*, 54 (11), 2008.

- X. Mestre and M.A. Lagunas, "Modified subspace algorithms for DoA estimation with large arrays", *IEEE Trans. SP*, 56 (2), 2008.
- R.R. Nadakuditi, A. Edelman, "Sample eigenvalue based detection of high-dimension signals in white noise using relatively few samples", *IEEE Trans. SP*, vol. 56, no. 7, pp. 2625-2638, 2008.
- B. Nadler, "Nonparametric Detection of Signals by Information Theoretic Criteria: Performance Analysis and an Improved Estimator", *IEEE Trans. SP*, 58 (5), 2010.
- L.A. Pastur, M. Shcherbina, "Eigenvalue Distribution of Large Random Matrices", Mathematical Surveys and Monographs, Providence: American Mathematical Society, 2011.
- G.T. Pham, P. Loubaton, P. Vallet, "Performance analysis of spatial smoothing schemes in the context of large arrays", IEEE Trans. on Signal Processing, vol. 64, no. 1, pp. 160-172, January 2016, also available on Arxiv (arXiv:1503.08196).

- G.T. Pham, P. Loubaton, "Applications of large empirical spatio-temporal covariance matrix in multipath channels detection", Signal Processing Conference (EUSIPCO), Nice, September 2015
- G.T. Pham, P. Loubaton, "Optimization of the loading factor of regularized estimated spatio-temporal Wiener filters in the large system case", IEEE SSP Workshop, Palma de Majorque, June 2016.
- P. Vallet, Ph. Loubaton, and X. Mestre, "Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case", *IEEE Trans. IT*, vol. 58, no. 2, pp. 1043-1068, 2012.
- P. Vallet, X. Mestre, P. Loubaton, "Performance analysis of an improved MUSIC DoA estimator", IEEE Trans. on Signal Processing, vol. 63, no. 23, pp. 6407-6422, December 1 2015, also available on Arxiv (arXiv:1503.01271).