Eigenspace estimation for source localization using large random matrices

Pascal Vallet[(1)]
Joint work with Philippe Loubaton[(1)] and Xavier Mestre[(2)]

[(1)] LabInfo IGM (CNRS-UMR 8049) / Université Paris-Est
[(2)] Centre Tecnologic de Telecomunicacions de Catalunya (CTTC) / Barcelona

Table of Contents

1 Introduction

2 Random matrix theory results

3 Consistent estimation of eigenspace

4 Numerical evaluations

- We will assume that K source signals are received by an antenna array of M elements, and $K < M$. At time n, we receive

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n,$$

with

  - $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ the $M \times K$ "steering vectors" matrix with $\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)$ linearly independent.
  - $\mathbf{s}_n = [s_{1,n}, \ldots, s_{n,K}]$ the vector of non-observable transmitted signals, assumed deterministic,
  - $\mathbf{v}_n$ a gaussian white noise (zero mean, covariance $\sigma^2 \mathbf{I}_M$).

- $\theta_1, \ldots, \theta_K$ are the parameters of interest of the K sources, it can be either frequencies, direction of arrival (DoA)...

- We will assume that K source signals are received by an antenna array of M elements, and $K < M$. At time n, we receive

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n,$$

with

  - $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ the $M \times K$ "steering vectors" matrix with $\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)$ linearly independent.
  - $\mathbf{s}_n = [s_{1,n}, \ldots, s_{n,K}]$ the vector of non-observable transmitted signals, assumed deterministic,
  - $\mathbf{v}_n$ a gaussian white noise (zero mean, covariance $\sigma^2 \mathbf{I}_M$).
- $\theta_1, \ldots, \theta_K$ are the parameters of interest of the K sources, it can be either frequencies, direction of arrival (DoA)...

- We will assume that K source signals are received by an antenna array of M elements, and $K < M$. At time n, we receive

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n,$$

with
  - $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ the $M \times K$ "steering vectors" matrix with $\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)$ linearly independent.
  - $\mathbf{s}_n = [s_{1,n}, \ldots, s_{n,K}]$ the vector of non-observable transmitted signals, assumed deterministic,
  - $\mathbf{v}_n$ a gaussian white noise (zero mean, covariance $\sigma^2 \mathbf{I}_M$).
- $\theta_1, \ldots, \theta_K$ are the parameters of interest of the K sources, it can be either frequencies, direction of arrival (DoA)...

- We will assume that K source signals are received by an antenna array of M elements, and $K < M$. At time n, we receive

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n,$$

  with
  - $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ the $M \times K$ "steering vectors" matrix with $\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)$ linearly independent.
  - $\mathbf{s}_n = [s_{1,n}, \ldots, s_{n,K}]$ the vector of non-observable transmitted signals, assumed deterministic,
  - $\mathbf{v}_n$ a gaussian white noise (zero mean, covariance $\sigma^2 \mathbf{I}_M$).
- $\theta_1, \ldots, \theta_K$ are the parameters of interest of the K sources, it can be either frequencies, direction of arrival (DoA)...

- We will assume that K source signals are received by an antenna array of M elements, and $K < M$. At time n, we receive

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n,$$

with

  - $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ the $M \times K$ "steering vectors" matrix with $\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)$ linearly independent.
  - $\mathbf{s}_n = [s_{1,n}, \ldots, s_{n,K}]$ the vector of non-observable transmitted signals, assumed deterministic,
  - $\mathbf{v}_n$ a gaussian white noise (zero mean, covariance $\sigma^2 \mathbf{I}_M$).

- $\theta_1, \ldots, \theta_K$ are the parameters of interest of the K sources, it can be either frequencies, direction of arrival (DoA)...

- We collect N observations of the previous model, stacked in $\mathbf{Y}_N = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$, and we can write

$$\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$$

with $\mathbf{S}_N$ and $\mathbf{V}_N$ built as $\mathbf{Y}_N$.

The goal is to infer the angles $\theta_1, \ldots, \theta_K$ from $\mathbf{Y}_N$.

- There are essentially two common methods:
  - Maximum Likelihood (ML) estimation
  - Subspace method.

- We collect N observations of the previous model, stacked in $\mathbf{Y}_N = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$, and we can write

$$\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$$

with $\mathbf{S}_N$ and $\mathbf{V}_N$ built as $\mathbf{Y}_N$.

The goal is to infer the angles $\theta_1, \ldots, \theta_K$ from $\mathbf{Y}_N$.

- There are essentially two common methods:
  - Maximum Likelihood (ML) estimation
  - Subspace method.

- We collect N observations of the previous model, stacked in $\mathbf{Y}_N = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$, and we can write

$$\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$$

with $\mathbf{S}_N$ and $\mathbf{V}_N$ built as $\mathbf{Y}_N$.

The goal is to infer the angles $\theta_1, \ldots, \theta_K$ from $\mathbf{Y}_N$.

- There are essentially two common methods:
  - Maximum Likelihood (ML) estimation
  - Subspace method.

- The ML estimator is given by

$$\underset{\boldsymbol{\omega}}{\operatorname{argmin}} \frac{1}{N} \operatorname{Tr}\left(\mathbf{I}_M - \mathbf{A}(\boldsymbol{\omega})(\mathbf{A}(\boldsymbol{\omega})^* \mathbf{A}(\boldsymbol{\omega}))^{-1} \mathbf{A}(\boldsymbol{\omega})^*\right) \mathbf{Y}_N \mathbf{Y}_N^*,$$

  where $\mathbf{A}(\boldsymbol{\omega})$ is the matrix in which we have replaced $[\theta_1, \ldots, \theta_K]$ by the variable $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_K]$.
- This estimator is consistent when $M, N \to \infty$, however, it clearly requires a multidimensional optimization.
- An alternative, requiring a monodimensional search, has been found through the subspace method.

- The ML estimator is given by

$$\underset{\boldsymbol{\omega}}{\operatorname{argmin}} \frac{1}{N} \operatorname{Tr}\left(\mathbf{I}_M - \mathbf{A}(\boldsymbol{\omega})(\mathbf{A}(\boldsymbol{\omega})^* \mathbf{A}(\boldsymbol{\omega}))^{-1} \mathbf{A}(\boldsymbol{\omega})^*\right) \mathbf{Y}_N \mathbf{Y}_N^*,$$

  where $\mathbf{A}(\boldsymbol{\omega})$ is the matrix in which we have replaced $[\theta_1, \ldots, \theta_K]$ by the variable $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_K]$.
- This estimator is consistent when $M, N \to \infty$, however, it clearly requires a multidimensional optimization.
- An alternative, requiring a monodimensional search, has been found through the subspace method.

- The ML estimator is given by

$$\underset{\boldsymbol{\omega}}{\operatorname{argmin}} \frac{1}{N} \operatorname{Tr}\left(\mathbf{I}_M - \mathbf{A}(\boldsymbol{\omega})(\mathbf{A}(\boldsymbol{\omega})^* \mathbf{A}(\boldsymbol{\omega}))^{-1} \mathbf{A}(\boldsymbol{\omega})^*\right) \mathbf{Y}_N \mathbf{Y}_N^*,$$

  where $\mathbf{A}(\boldsymbol{\omega})$ is the matrix in which we have replaced $[\theta_1, \ldots, \theta_K]$ by the variable $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_K]$.
- This estimator is consistent when $M, N \to \infty$, however, it clearly requires a multidimensional optimization.
- An alternative, requiring a monodimensional search, has been found through the subspace method.

- Assuming $\mathbf{S}_N$ has full rank K, then $\frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^*$ has K non null eigenvalues

$$0 = \lambda_{1,N} = \ldots = \lambda_{M-K,N} < \lambda_{M-K+1,N} < \ldots < \lambda_{M,N}.$$

  We denote by $\mathbf{\Pi}_N$ the projector onto the eigensubspace associated with eigenvalue 0.

- Since span$\{\mathbf{a}(\theta_1),\ldots,\mathbf{a}(\theta_K)\}$ is also the eigenspace associated with non null eigenvalues $\lambda_{M-K+1,N},\ldots,\lambda_{M,N}$, it is possible to determine the $(\theta_k)_{k=1,\ldots,K}$.

MUSIC algorithm

The angles $\theta_1,\ldots,\theta_K$ are the (unique) solutions of the equation

$$\eta(\theta) := \mathbf{a}(\theta)^*\mathbf{\Pi}_N\mathbf{a}(\theta) = 0.$$

- Assuming $\mathbf{S}_N$ has full rank K, then $\frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^*$ has K non null eigenvalues

$$0 = \lambda_{1,N} = \ldots = \lambda_{M-K,N} < \lambda_{M-K+1,N} < \ldots < \lambda_{M,N}.$$

  We denote by $\mathbf{\Pi}_N$ the projector onto the eigensubspace associated with eigenvalue 0.

- Since span$\{\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)\}$ is also the eigenspace associated with non null eigenvalues $\lambda_{M-K+1,N}, \ldots, \lambda_{M,N}$, it is possible to determine the $(\theta_k)_{k=1,\ldots,K}$.

MUSIC algorithm

The angles $\theta_1, \ldots, \theta_K$ are the (unique) solutions of the equation

$$\eta(\theta) := \mathbf{a}(\theta)^* \mathbf{\Pi}_N \mathbf{a}(\theta) = 0.$$

- Assuming $\mathbf{S}_N$ has full rank K, then $\frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^*$ has K non null eigenvalues

$$0 = \lambda_{1,N} = \ldots = \lambda_{M-K,N} < \lambda_{M-K+1,N} < \ldots < \lambda_{M,N}.$$

  We denote by $\mathbf{\Pi}_N$ the projector onto the eigensubspace associated with eigenvalue 0.

- Since span$\{\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)\}$ is also the eigenspace associated with non null eigenvalues $\lambda_{M-K+1,N}, \ldots, \lambda_{M,N}$, it is possible to determine the $(\theta_k)_{k=1,\ldots,K}$.

MUSIC algorithm

The angles $\theta_1, \ldots, \theta_K$ are the (unique) solutions of the equation

$$\eta(\theta) := \mathbf{a}(\theta)^* \mathbf{\Pi}_N \mathbf{a}(\theta) = 0.$$

- We denote by $\hat{\lambda}_{1,N} \leq \ldots \leq \hat{\lambda}_{M,N}$ the eigenvalues of $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$, and $\hat{\mathbf{u}}_{1,N}, \ldots, \hat{\mathbf{u}}_{M,N}$ the associated eigenvectors.
- In practice, to estimate the angles, we only know $\mathbf{Y}_N$, and we estimate function $\eta(\theta)$ by

$$\hat{\eta}_{trad}(\theta) := \mathbf{a}(\theta)^* \hat{\mathbf{\Pi}}_N \mathbf{a}(\theta),$$

  with $\hat{\mathbf{\Pi}}_N = \sum_{k=1}^{M-K} \hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*$ the projector onto the eigensubspace associated with $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N}$.
- In the case where $N \to \infty$ while M is constant, this estimator is consistent because $\|\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^* - \frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^*\| \to 0$ a.s.
- However, when $M, N \to \infty$ while $c_N = M/N \to c > 0$, the previous convergence fails and the estimator is no more consistent.

- We denote by $\hat{\lambda}_{1,N} \leq \ldots \leq \hat{\lambda}_{M,N}$ the eigenvalues of $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$, and $\hat{\mathbf{u}}_{1,N}, \ldots, \hat{\mathbf{u}}_{M,N}$ the associated eigenvectors.
- In practice, to estimate the angles, we only know $\mathbf{Y}_N$, and we estimate function $\eta(\theta)$ by

$$\hat{\eta}_{trad}(\theta) := \mathbf{a}(\theta)^* \hat{\mathbf{\Pi}}_N \mathbf{a}(\theta),$$

with $\hat{\mathbf{\Pi}}_N = \sum_{k=1}^{M-K} \hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*$ the projector onto the eigensubspace associated with $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N}$.

- In the case where $N \to \infty$ while M is constant, this estimator is consistent because $\|\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^* - \frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^*\| \to 0$ a.s.
- However, when $M, N \to \infty$ while $c_N = M/N \to c > 0$, the previous convergence fails and the estimator is no more consistent.

- We denote by $\hat{\lambda}_{1,N} \leq \ldots \leq \hat{\lambda}_{M,N}$ the eigenvalues of $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$, and $\hat{\mathbf{u}}_{1,N}, \ldots, \hat{\mathbf{u}}_{M,N}$ the associated eigenvectors.
- In practice, to estimate the angles, we only know $\mathbf{Y}_N$, and we estimate function $\eta(\theta)$ by

$$\hat{\eta}_{trad}(\theta) := \mathbf{a}(\theta)^* \hat{\mathbf{\Pi}}_N \mathbf{a}(\theta),$$

with $\hat{\mathbf{\Pi}}_N = \sum_{k=1}^{M-K} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^*$ the projector onto the eigensubspace associated with $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N}$.

- In the case where $N \to \infty$ while M is constant, this estimator is consistent because $\| \frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^* - \frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^* \| \to 0$ a.s.
- However, when $M, N \to \infty$ while $c_N = M/N \to c > 0$, the previous convergence fails and the estimator is no more consistent.

- We denote by $\hat{\lambda}_{1,N} \leq \ldots \leq \hat{\lambda}_{M,N}$ the eigenvalues of $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$, and $\hat{\mathbf{u}}_{1,N}, \ldots, \hat{\mathbf{u}}_{M,N}$ the associated eigenvectors.
- In practice, to estimate the angles, we only know $\mathbf{Y}_N$, and we estimate function $\eta(\theta)$ by

$$\hat{\eta}_{trad}(\theta) := \mathbf{a}(\theta)^* \hat{\mathbf{\Pi}}_N \mathbf{a}(\theta),$$

  with $\hat{\mathbf{\Pi}}_N = \sum_{k=1}^{M-K} \hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*$ the projector onto the eigensubspace associated with $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N}$.
- In the case where $N \to \infty$ while M is constant, this estimator is consistent because $\| \frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^* - \frac{1}{N}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^* \| \to 0$ a.s.
- However, when $M, N \to \infty$ while $c_N = M/N \to c > 0$, the previous convergence fails and the estimator is no more consistent.

- For convenience of notations, we rewrite the main model

$$\mathbf{\Sigma}_N := \frac{\mathbf{Y}_N}{\sqrt{N}}, \quad \mathbf{B}_N := \frac{\mathbf{A}\mathbf{S}_N}{\sqrt{N}}, \quad \mathbf{W}_N := \frac{\mathbf{V}_N}{\sqrt{N}},$$

so that $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ is the well-known gaussian information plus noise model.

Problem

Find a consistent estimator of the quadratic form $\mathbf{d}_N^* \mathbf{\Pi}_N \mathbf{d}_N$ in the case where

- $\sup_N \|\mathbf{B}_N\| < \infty$,
- $\sup_N \|\mathbf{d}_N\| < \infty$,
- $\mathbf{W}_N$ gaussian i.i.d (zero mean and elements having variance σ^2/N),
- $M, N \to \infty$ while $c_N = M/N \to c \in (0,1)$.

- For convenience of notations, we rewrite the main model

$$\mathbf{\Sigma}_N := \frac{\mathbf{Y}_N}{\sqrt{N}}, \quad \mathbf{B}_N := \frac{\mathbf{A}\mathbf{S}_N}{\sqrt{N}}, \quad \mathbf{W}_N := \frac{\mathbf{V}_N}{\sqrt{N}},$$

so that $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ is the well-known gaussian information plus noise model.

Problem

Find a consistent estimator of the quadratic form $\mathbf{d}_N^* \mathbf{\Pi}_N \mathbf{d}_N$ in the case where

- $\sup_N \|\mathbf{B}_N\| < \infty$,
- $\sup_N \|\mathbf{d}_N\| < \infty$,
- $\mathbf{W}_N$ gaussian i.i.d (zero mean and elements having variance σ^2/N),
- $M, N \to \infty$ while $c_N = M/N \to c \in (0,1)$.

- Some related works
  - Mestre (2008) derived an estimator of the previous quadratic form, in the case where the source signals matrix $\mathbf{S}_N$ is gaussian i.i.d. In this case, $\mathbf{\Sigma}_N$ has the same distribution as $(\mathbf{A}\mathbf{A}^* + \sigma^2 \mathbf{I}_M)\mathbf{X}_N$ with $\mathbf{X}_N$ gaussian i.i.d.
  - Couillet et al. (2010) extend this work to the case where $\mathbf{S}_N$ is i.i.d but not necessarily gaussian.
- For the remainder of the talk, we define some shortcuts
  - $N \to \infty$ stands for the previous regime of convergence $M, N \to \infty$ while $c_N = M/N \to c \in (0,1)$.
  - For two sequences of r.v $(X_N), (Y_N)$, $X_N \asymp Y_N$ for $X_N - Y_N \to 0$ a.s as $N \to \infty$

- Some related works
  - Mestre (2008) derived an estimator of the previous quadratic form, in the case where the source signals matrix $\mathbf{S}_N$ is gaussian i.i.d. In this case, $\mathbf{\Sigma}_N$ has the same distribution as $(\mathbf{A}\mathbf{A}^* + \sigma^2 \mathbf{I}_M)\mathbf{X}_N$ with $\mathbf{X}_N$ gaussian i.i.d.
  - Couillet et al. (2010) extend this work to the case where $\mathbf{S}_N$ is i.i.d but not necessarily gaussian.
- For the remainder of the talk, we define some shortcuts
  - $N \to \infty$ stands for the previous regime of convergence $M, N \to \infty$ while $c_N = M/N \to c \in (0,1)$.
  - For two sequences of r.v $(X_N), (Y_N)$, $X_N \asymp Y_N$ for $X_N - Y_N \to 0$ a.s as $N \to \infty$

- Some related works
  - Mestre (2008) derived an estimator of the previous quadratic form, in the case where the source signals matrix $\mathbf{S}_N$ is gaussian i.i.d. In this case, $\mathbf{\Sigma}_N$ has the same distribution as $(\mathbf{A}\mathbf{A}^* + \sigma^2 \mathbf{I}_M)\mathbf{X}_N$ with $\mathbf{X}_N$ gaussian i.i.d.
  - Couillet et al. (2010) extend this work to the case where $\mathbf{S}_N$ is i.i.d but not necessarily gaussian.
- For the remainder of the talk, we define some shortcuts
  - $N \to \infty$ stands for the previous regime of convergence $M, N \to \infty$ while $c_N = M/N \to c \in (0,1)$.
  - For two sequences of r.v $(X_N), (Y_N)$, $X_N \asymp Y_N$ for $X_N - Y_N \to 0$ a.s as $N \to \infty$

- Some related works
  - Mestre (2008) derived an estimator of the previous quadratic form, in the case where the source signals matrix $\mathbf{S}_N$ is gaussian i.i.d. In this case, $\mathbf{\Sigma}_N$ has the same distribution as $(\mathbf{A}\mathbf{A}^* + \sigma^2 \mathbf{I}_M)\mathbf{X}_N$ with $\mathbf{X}_N$ gaussian i.i.d.
  - Couillet et al. (2010) extend this work to the case where $\mathbf{S}_N$ is i.i.d but not necessarily gaussian.
- For the remainder of the talk, we define some shortcuts
  - $N \to \infty$ stands for the previous regime of convergence $M, N \to \infty$ while $c_N = M/N \to c \in (0, 1)$.
  - For two sequences of r.v $(X_N), (Y_N)$, $X_N \asymp Y_N$ for $X_N - Y_N \to 0$ a.s as $N \to \infty$

- Some related works
  - Mestre (2008) derived an estimator of the previous quadratic form, in the case where the source signals matrix $\mathbf{S}_N$ is gaussian i.i.d. In this case, $\mathbf{\Sigma}_N$ has the same distribution as $(\mathbf{A}\mathbf{A}^* + \sigma^2 \mathbf{I}_M)\mathbf{X}_N$ with $\mathbf{X}_N$ gaussian i.i.d.
  - Couillet et al. (2010) extend this work to the case where $\mathbf{S}_N$ is i.i.d but not necessarily gaussian.
- For the remainder of the talk, we define some shortcuts
  - $N \to \infty$ stands for the previous regime of convergence $M, N \to \infty$ while $c_N = M/N \to c \in (0,1)$.
  - For two sequences of r.v $(X_N), (Y_N)$, $X_N \asymp Y_N$ for $X_N - Y_N \to 0$ a.s as $N \to \infty$

Table of Contents

1 Introduction

2 Random matrix theory results

3 Consistent estimation of eigenspace

4 Numerical evaluations

- Let $\hat{\mu}_N = \frac{1}{M}\sum_{k=1}^{M}\delta_{\hat{\lambda}_{k,N}}$ the e.s.d of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$, and its Stieltjes transform

$$\hat{m}_N(z) := \int_{\mathbb{R}^+} \frac{\mathrm{d}\hat{\mu}_N(\lambda)}{\lambda - z} := \frac{1}{M}\mathrm{Tr}(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1} \text{ for } z \in \mathbb{C}\backslash\mathbb{R}^+.$$

Theorem (Dozier-Silverstein (2007))

As $N \to \infty$, $\hat{m}_N(z) \asymp m_N(z)$ with $m_N(z)$ the Stieltjes transform of a deterministic distribution μ_N, and unique solution to the equation $m_N(z) := \frac{1}{M}\mathrm{Tr}\,\mathbf{T}_N(z)$ with

$$\mathbf{T}_N(z) := \left(\frac{\mathbf{B}_N\mathbf{B}_N^*}{1+\sigma^2 c_N m_N(z)} - z(1+\sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1-c_N)\mathbf{I}_M\right)^{-1}.$$

- The same result holds for quadratic form of the resolvent (Hachem et al.(2010)), for $\mathbf{d}_N$ uniformly bounded in N,

$$\mathbf{d}_N^*(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}\mathbf{d}_N \asymp \mathbf{d}_N^*\mathbf{T}_N(z)\mathbf{d}_N.$$

- Let $\hat{\mu}_N = \frac{1}{M}\sum_{k=1}^{M} \delta_{\hat{\lambda}_{k,N}}$ the e.s.d of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$, and its Stieltjes transform

$$\hat{m}_N(z) := \int_{\mathbb{R}^+} \frac{d\hat{\mu}_N(\lambda)}{\lambda - z} := \frac{1}{M} \mathrm{Tr}(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1} \text{ for } z \in \mathbb{C} \backslash \mathbb{R}^+.$$

Theorem (Dozier-Silverstein (2007))

As $N \to \infty$, $\hat{m}_N(z) \asymp m_N(z)$ with $m_N(z)$ the Stieltjes transform of a deterministic distribution μ_N, and unique solution to the equation $m_N(z) := \frac{1}{M} \mathrm{Tr}\, \mathbf{T}_N(z)$ with

$$\mathbf{T}_N(z) := \left(\frac{\mathbf{B}_N \mathbf{B}_N^*}{1 + \sigma^2 c_N m_N(z)} - z(1 + \sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2 (1 - c_N)\mathbf{I}_M \right)^{-1}.$$

- The same result holds for quadratic form of the resolvent (Hachem et al.(2010)), for $\mathbf{d}_N$ uniformly bounded in N,

$$\mathbf{d}_N^* (\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1} \mathbf{d}_N \asymp \mathbf{d}_N^* \mathbf{T}_N(z) \mathbf{d}_N.$$

- Let $\hat{\mu}_N = \frac{1}{M}\sum_{k=1}^{M} \delta_{\hat{\lambda}_{k,N}}$ the e.s.d of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$, and its Stieltjes transform

$$\hat{m}_N(z) := \int_{\mathbb{R}^+} \frac{\mathrm{d}\hat{\mu}_N(\lambda)}{\lambda - z} := \frac{1}{M}\mathrm{Tr}(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1} \text{ for } z \in \mathbb{C}\backslash\mathbb{R}^+.$$

Theorem (Dozier-Silverstein (2007))

As $N \to \infty$, $\hat{m}_N(z) \asymp m_N(z)$ with $m_N(z)$ the Stieltjes transform of a deterministic distribution μ_N, and unique solution to the equation $m_N(z) := \frac{1}{M}\mathrm{Tr}\,\mathbf{T}_N(z)$ with

$$\mathbf{T}_N(z) := \left(\frac{\mathbf{B}_N\mathbf{B}_N^*}{1+\sigma^2 c_N m_N(z)} - z(1+\sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1-c_N)\mathbf{I}_M\right)^{-1}.$$

- The same result holds for quadratic form of the resolvent (Hachem et al.(2010)), for $\mathbf{d}_N$ uniformly bounded in N,

$$\mathbf{d}_N^*(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}\mathbf{d}_N \asymp \mathbf{d}_N^*\mathbf{T}_N(z)\mathbf{d}_N.$$

- The following result is a rephrasing of the result of Dozier-Silverstein (2007) about the support of μ_N.
- Let $f_N(w) = \frac{1}{M}\operatorname{Tr}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1}$ and

$$\phi_N(w) = w\left(1-\sigma^2 c_N f_N(w)\right)^2 + \sigma^2(1-c)\left(1-\sigma^2 c_N f_N(w)\right).$$

Theorem

The support $\operatorname{supp}(\mu_N)$ *is the union of* $1 \le Q \le K+1$ *compact intervals*

$$\operatorname{supp}(\mu_N) = \bigcup_{q=1}^{Q} [x_{q,N}^-, x_{q,N}^+],$$

with $\{x_{q,N}^-, x_{q,N}^+\}_{q=1,\dots,Q}$ *the positive local extrema of* ϕ_N *and* $x_{1,N}^- > 0$.

- The following result is a rephrasing of the result of Dozier-Silverstein (2007) about the support of μ_N.
- Let $f_N(w) = \frac{1}{M}\operatorname{Tr}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1}$ and

$$\phi_N(w) = w\left(1 - \sigma^2 c_N f_N(w)\right)^2 + \sigma^2(1-c)\left(1 - \sigma^2 c_N f_N(w)\right).$$

Theorem

The support $\operatorname{supp}(\mu_N)$ *is the union of* $1 \le Q \le K+1$ *compact intervals*

$$\operatorname{supp}(\mu_N) = \bigcup_{q=1}^{Q} [x_{q,N}^-, x_{q,N}^+],$$

with $\{x_{q,N}^-, x_{q,N}^+\}_{q=1,\ldots,Q}$ *the positive local extrema of* ϕ_N *and* $x_{1,N}^- > 0$.

- The following result is a rephrasing of the result of Dozier-Silverstein (2007) about the support of μ_N.
- Let $f_N(w) = \frac{1}{M}\operatorname{Tr}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1}$ and

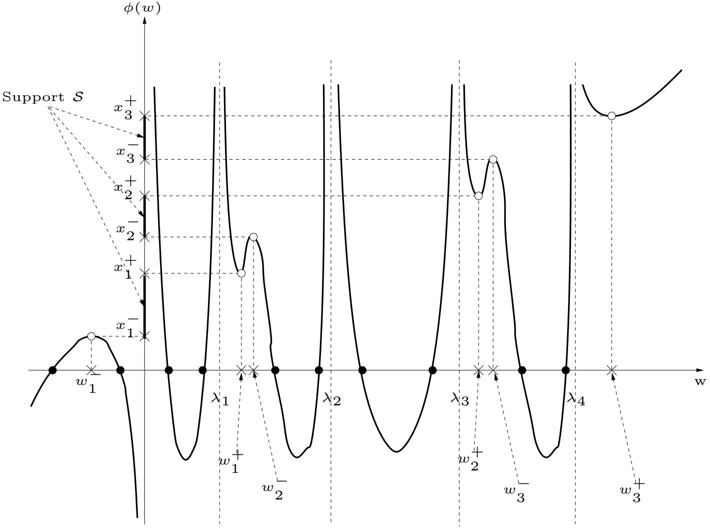
$$\phi_N(w) = w\left(1 - \sigma^2 c_N f_N(w)\right)^2 + \sigma^2(1-c)\left(1 - \sigma^2 c_N f_N(w)\right).$$

Theorem

The support $\operatorname{supp}(\mu_N)$ *is the union of* $1 \leq Q \leq K+1$ *compact intervals*

$$\operatorname{supp}(\mu_N) = \bigcup_{q=1}^{Q} [x_{q,N}^-, x_{q,N}^+],$$

with $\{x_{q,N}^-, x_{q,N}^+\}_{q=1,\ldots,Q}$ *the positive local extrema of* ϕ_N *and* $x_{1,N}^- > 0$.


$\phi(w)$
Support $\mathcal{S}$
x_3^+
x_3^-
x_2^+
x_2^-
x_1^+
x_1^-
w_1^-
λ_1
λ_2
λ_3
λ_4
w
w_1^+
w_2^-
w_2^+
w_3^-
w_3^+

- Each eigenvalue $0, \lambda_{1,N}, \ldots, \lambda_{M,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ belongs to an interval $]w_{q,N}^-, w_{q,N}^+[$.
- An eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ is said to be associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$ if $\lambda_{k,N} \in]w_{q,N}^-, w_{q,N}^+[$.
- Two eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are "separated" if they are associated with two different clusters.
- If the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are sufficiently spaced, σ and/or c_N are small enough, all the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are separated, i.e we have exactly $Q = K + 1$ disjoint compact intervals in the support of μ_N.

- Each eigenvalue $0, \lambda_{1,N}, \ldots, \lambda_{M,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ belongs to an interval $]w_{q,N}^-, w_{q,N}^+[$.
- An eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ is said to be associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$ if $\lambda_{k,N} \in]w_{q,N}^-, w_{q,N}^+[$.
- Two eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are "separated" if they are associated with two different clusters.
- If the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are sufficiently spaced, σ and/or c_N are small enough, all the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are separated, i.e we have exactly $Q = K + 1$ disjoint compact intervals in the support of μ_N.

- Each eigenvalue $0, \lambda_{1,N}, \ldots, \lambda_{M,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ belongs to an interval $]w_{q,N}^-, w_{q,N}^+[$.
- An eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ is said to be associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$ if $\lambda_{k,N} \in]w_{q,N}^-, w_{q,N}^+[$.
- Two eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are "separated" if they are associated with two different clusters.
- If the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are sufficiently spaced, σ and/or c_N are small enough, all the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ are separated, i.e we have exactly $Q = K + 1$ disjoint compact intervals in the support of μ_N.

- Each eigenvalue $0, \lambda_{1,N}, \ldots, \lambda_{M,N}$ of $\mathbf{B}_N \mathbf{B}_N^*$ belongs to an interval $]w_{q,N}^-, w_{q,N}^+[$.
- An eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N \mathbf{B}_N^*$ is said to be associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$ if $\lambda_{k,N} \in]w_{q,N}^-, w_{q,N}^+[$.
- Two eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ are "separated" if they are associated with two different clusters.
- If the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ are sufficiently spaced, σ and/or c_N are small enough, all the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ are separated, i.e we have exactly $Q = K + 1$ disjoint compact intervals in the support of μ_N.

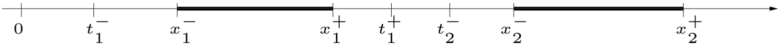
Theorem

Assume

- 0 *is separated from the other eigenvalues, i.e* 0 *is the unique eigenvalue associated with* $[x^-_{1,N}, x^+_{1,N}]$,
- $\exists\ t^-_1, t^+_1, t^-_2$ *independent of* N *s.t* $0 < t^-_1 < \inf_N x^-_{1,N}$ *and* $t^-_2 > t^+_1 > \sup_N x^+_{1,N}$,

then, for all large N*, w.p.1,*

$$\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N} \in]t^-_1, t^+_1[\quad \textit{and} \quad \hat{\lambda}_{M-K+1,N} > t^-_2.$$



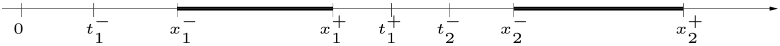
Theorem

Assume

- 0 *is separated from the other eigenvalues, i.e* 0 *is the unique eigenvalue associated with* $[x_{1,N}^-, x_{1,N}^+]$,
- $\exists\, t_1^-, t_1^+, t_2^-$ *independent of* N *s.t* $0 < t_1^- < \inf_N x_{1,N}^-$ *and* $t_2^- > t_1^+ > \sup_N x_{1,N}^+$,

then, for all large N, w.p.1,

$$\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N} \in]t_1^-, t_1^+[\quad \textit{and} \quad \hat{\lambda}_{M-K+1,N} > t_2^-.$$



Theorem

Assume

- 0 *is separated from the other eigenvalues, i.e* 0 *is the unique eigenvalue associated with* $[x_{1,N}^-, x_{1,N}^+]$,
- $\exists\ t_1^-, t_1^+, t_2^-$ *independent of* N *s.t* $0 < t_1^- < \inf_N x_{1,N}^-$ *and* $t_2^- > t_1^+ > \sup_N x_{1,N}^+$,

then, for all large N, w.p.1,

$$\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N} \in]t_1^-, t_1^+[\quad \text{and} \quad \hat{\lambda}_{M-K+1,N} > t_2^-.$$

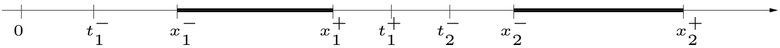


Table of Contents

1 Introduction

2 Random matrix theory results

3 Consistent estimation of eigenspace

4 Numerical evaluations

- We want to estimate the quadratic form $\eta_N := \mathbf{d}_N^* \mathbf{\Pi}_N \mathbf{d}_N$ of the noise subspace projector, **under the assumption that** 0 **is the unique eigenvalue associated to** $[x_{1,N}^-, x_{1,N}^+]$ **for all large** N.
- **No assumption is made on the number of sources** K **which may scale-up with** N **or stay constant.**
- From residues theorem, we get

$$\eta_N = \frac{1}{2\pi \mathbf{i}} \oint_{\mathscr{C}^-} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - \lambda \mathbf{I}_M \right)^{-1} \mathbf{d}_N \mathrm{d}\lambda,$$

  with $\mathscr{C}^-$ a clockwise oriented closed path enclosing 0 and no other eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$.
- The fundamental point is that we can find such a path which can be parametrized by a function of m_N, the Stieltjes transform of μ_N.

- We want to estimate the quadratic form $\eta_N := \mathbf{d}_N^* \mathbf{\Pi}_N \mathbf{d}_N$ of the noise subspace projector, **under the assumption that** 0 **is the unique eigenvalue associated to** $[x_{1,N}^-, x_{1,N}^+]$ **for all large** N.
- **No assumption is made on the number of sources** K **which may scale-up with** N **or stay constant.**
- From residues theorem, we get

$$\eta_N = \frac{1}{2\pi\mathbf{i}} \oint_{\mathscr{C}^-} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - \lambda \mathbf{I}_M\right)^{-1} \mathbf{d}_N \mathrm{d}\lambda,$$

  with $\mathscr{C}^-$ a clockwise oriented closed path enclosing 0 and no other eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$.
- The fundamental point is that we can find such a path which can be parametrized by a function of m_N, the Stieltjes transform of μ_N.

- We want to estimate the quadratic form $\eta_N := \mathbf{d}_N^* \mathbf{\Pi}_N \mathbf{d}_N$ of the noise subspace projector, **under the assumption that** 0 **is the unique eigenvalue associated to** $[x_{1,N}^-, x_{1,N}^+]$ **for all large** N.
- **No assumption is made on the number of sources** K **which may scale-up with** N **or stay constant.**
- From residues theorem, we get

$$\eta_N = \frac{1}{2\pi\mathbf{i}} \oint_{\mathscr{C}^-} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - \lambda \mathbf{I}_M\right)^{-1} \mathbf{d}_N \mathrm{d}\lambda,$$

  with $\mathscr{C}^-$ a clockwise oriented closed path enclosing 0 and no other eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$.
- The fundamental point is that we can find such a path which can be parametrized by a function of m_N, the Stieltjes transform of μ_N.

- We want to estimate the quadratic form $\eta_N := \mathbf{d}_N^* \mathbf{\Pi}_N \mathbf{d}_N$ of the noise subspace projector, **under the assumption that** 0 **is the unique eigenvalue associated to** $[x_{1,N}^-, x_{1,N}^+]$ **for all large** N.
- **No assumption is made on the number of sources** K **which may scale-up with** N **or stay constant.**
- From residues theorem, we get

$$\eta_N = \frac{1}{2\pi \mathbf{i}} \oint_{\mathscr{C}^-} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - \lambda \mathbf{I}_M \right)^{-1} \mathbf{d}_N \mathrm{d}\lambda,$$

  with $\mathscr{C}^-$ a clockwise oriented closed path enclosing 0 and no other eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$.
- The fundamental point is that we can find such a path which can be parametrized by a function of m_N, the Stieltjes transform of μ_N.

- Consider the function

$$w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2 c_N(1+\sigma^2 c_N m_N(z)) \quad z \in \mathbb{C}\setminus\mathbb{R}^+.$$

- The following limit exists (Dozier-Silverstein (2007)), for $x \in \mathbb{R}$,

$$w_N(x) := \lim_{\substack{z \to x \\ \mathrm{Im}\{z\}>0}} w_N(z).$$

- We consider $\mathcal{C} = \left\{w_N(x) : x \in [t_1^-, t_1^+]\right\} \cup \left\{w_N(x)^* : x \in [t_1^-, t_1^+]\right\}$.

- Consider the function

$$w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2 c_N(1+\sigma^2 c_N m_N(z)) \quad z \in \mathbb{C} \backslash \mathbb{R}^+.$$

- The following limit exists (Dozier-Silverstein (2007)), for $x \in \mathbb{R}$,

$$w_N(x) := \lim_{\substack{z \to x \\ \mathrm{Im}\{z\} > 0}} w_N(z).$$

- We consider $\mathscr{C} = \left\{ w_N(x) : x \in [t_1^-, t_1^+] \right\} \cup \left\{ w_N(x)^* : x \in [t_1^-, t_1^+] \right\}$.

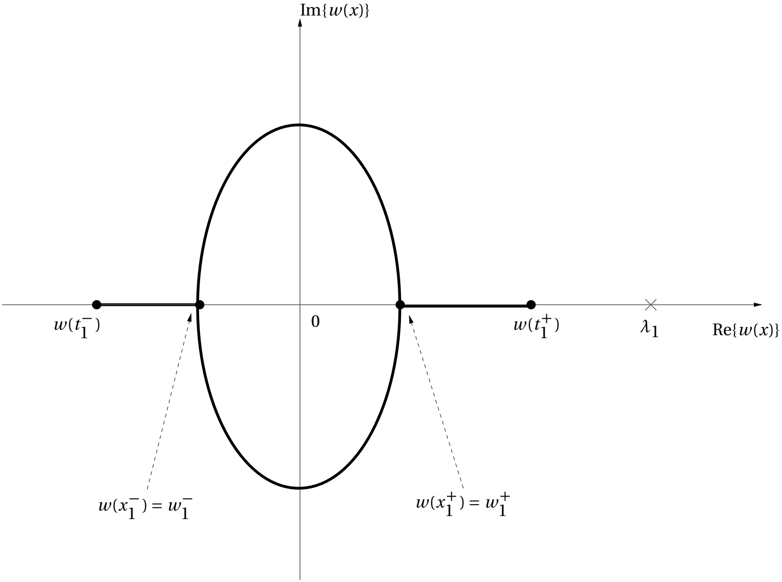
- Consider the function

$$w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2 c_N (1+\sigma^2 c_N m_N(z)) \quad z \in \mathbb{C}\backslash\mathbb{R}^+.$$

- The following limit exists (Dozier-Silverstein (2007)), for $x \in \mathbb{R}$,

$$w_N(x) := \lim_{\substack{z \to x \\ \mathrm{Im}\{z\}>0}} w_N(z).$$

- We consider $\mathscr{C} = \left\{ w_N(x) : x \in [t_1^-, t_1^+] \right\} \cup \left\{ w_N(x)^* : x \in [t_1^-, t_1^+] \right\}$.


Im{$w(x)$}
$w(t_1^-)$
0
$w(t_1^+)$
λ_1
Re{$w(x)$}
$w(x_1^-) = w_1^-$
$w(x_1^+) = w_1^+$

- This allows to rewrite the previous integral as

$$\eta_N = \frac{1}{\pi}\mathrm{Im}\left\{\int_{t_1^-}^{t_1^+} \mathbf{d}_N^* \left(\mathbf{B}_N\mathbf{B}_N^* - w_N(x)\mathbf{I}_M\right)^{-1} \mathbf{d}_N w_N'(x)\mathrm{d}x\right\}.$$

- Dominated convergence can be applied to obtain

$$\eta_N = \frac{1}{\pi}\lim_{y\downarrow 0}\mathrm{Im}\left\{\int_{t_1^-}^{t_1^+} \mathbf{d}_N^* \left(\mathbf{B}_N\mathbf{B}_N^* - w_N(x+\mathbf{i}y)\mathbf{I}_M\right)^{-1} \mathbf{d}_N w_N'(x+\mathbf{i}y)\mathrm{d}x\right\}$$

$$= \lim_{y\downarrow 0}\frac{1}{2\pi\mathbf{i}}\oint_{\partial\mathscr{R}_y^-} \mathbf{d}_N^* \left(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M\right)^{-1} \mathbf{d}_N w_N'(z)\mathrm{d}z,$$

with, for $y > 0$, $\partial\mathscr{R}_y^-$ the boundary of the rectangle

$$\mathscr{R}_y = \left\{u + \mathbf{i}v : u \in [t_1^-, t_1^+], v \in [-y, y]\right\}.$$

- The previous limit can be dropped, due to the holomorphy of $\mathbf{d}_N^* \left(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M\right)^{-1} \mathbf{d}_N w_N'(z)$ on $\mathbb{C}\setminus\mathrm{supp}(\mu_N)$.

- This allows to rewrite the previous integral as

$$\eta_N = \frac{1}{\pi} \mathrm{Im} \left\{ \int_{t_1^-}^{t_1^+} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(x) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(x) \mathrm{d}x \right\}.$$

- Dominated convergence can be applied to obtain

$$\begin{aligned} \eta_N &= \frac{1}{\pi} \lim_{y \downarrow 0} \mathrm{Im} \left\{ \int_{t_1^-}^{t_1^+} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(x + \mathbf{i}y) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(x + \mathbf{i}y) \mathrm{d}x \right\} \\ &= \lim_{y \downarrow 0} \frac{1}{2\pi \mathbf{i}} \oint_{\partial \mathscr{R}_y^-} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(z) \mathrm{d}z, \end{aligned}$$

with, for $y > 0$, $\partial \mathscr{R}_y^-$ the boundary of the rectangle

$$\mathscr{R}_y = \left\{ u + \mathbf{i}v : u \in [t_1^-, t_1^+], v \in [-y, y] \right\}.$$

- The previous limit can be dropped, due to the holomorphy of $\mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(z)$ on $\mathbb{C} \setminus \mathrm{supp}(\mu_N)$.

- This allows to rewrite the previous integral as

$$\eta_N = \frac{1}{\pi} \mathrm{Im} \left\{ \int_{t_1^-}^{t_1^+} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(x) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(x) \mathrm{d}x \right\}.$$

- Dominated convergence can be applied to obtain

$$\begin{aligned} \eta_N &= \frac{1}{\pi} \lim_{y \downarrow 0} \mathrm{Im} \left\{ \int_{t_1^-}^{t_1^+} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(x + \mathbf{i}y) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(x + \mathbf{i}y) \mathrm{d}x \right\} \\ &= \lim_{y \downarrow 0} \frac{1}{2\pi \mathbf{i}} \oint_{\partial \mathscr{R}_y^-} \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(z) \mathrm{d}z, \end{aligned}$$

  with, for $y > 0$, $\partial \mathscr{R}_y^-$ the boundary of the rectangle

$$\mathscr{R}_y = \left\{ u + \mathbf{i}v : u \in [t_1^-, t_1^+], v \in [-y, y] \right\}.$$

- The previous limit can be dropped, due to the holomorphy of $\mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(z)$ on $\mathbb{C} \setminus \mathrm{supp}(\mu_N)$.

- The previous integrand can be written as

$$g_N(z) := \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z) \mathbf{I}_M \right)^{-1} \mathbf{d}_N w_N'(z)$$
$$= \mathbf{d}_N^* \mathbf{T}_N(z) \mathbf{d}_N \frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)}.$$

- From the previous result, we have the following convergence

$$m_N(z) \asymp \hat{m}_N(z) = \frac{1}{M} \operatorname{Tr} \mathbf{Q}_N(z) \quad \text{and} \quad \mathbf{d}_N^* \mathbf{T}_N(z) \mathbf{d}_N \asymp \mathbf{d}_N^* \mathbf{Q}_N(z) \mathbf{d}_N,$$

with $\mathbf{Q}_N(z) = (\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* - z \mathbf{I}_M)^{-1}$.

- Let $\hat{g}_N(z) := \mathbf{d}_N^* \mathbf{Q}_N(z) \mathbf{d}_N \frac{\hat{w}_N'(z)}{1+\sigma^2 c_N \hat{m}_N(z)}$. We can show that

$$\left| \frac{1}{2\pi \mathbf{i}} \oint_{\partial \mathscr{R}_y^-} \left(g_N(z) - \hat{g}_N(z) \right) \mathrm{d}z \right| \asymp 0,$$

with $\hat{w}_N(z) = z(1 + \sigma^2 c_N \hat{m}_N(z))^2 - \sigma^2 c_N (1 + \sigma^2 c_N \hat{m}_N(z))$.

- The previous integrand can be written as

$$g_N(z) := \mathbf{d}_N^* \left(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M\right)^{-1} \mathbf{d}_N w_N'(z)$$
$$= \mathbf{d}_N^*\mathbf{T}_N(z)\mathbf{d}_N \frac{w_N'(z)}{1+\sigma^2 c_N m_N(z)}.$$

- From the previous result, we have the following convergence

$$m_N(z) \asymp \hat{m}_N(z) = \frac{1}{M}\operatorname{Tr}\mathbf{Q}_N(z) \quad \text{and} \quad \mathbf{d}_N^*\mathbf{T}_N(z)\mathbf{d}_N \asymp \mathbf{d}_N^*\mathbf{Q}_N(z)\mathbf{d}_N,$$

with $\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}$.

- Let $\hat{g}_N(z) := \mathbf{d}_N^*\mathbf{Q}_N(z)\mathbf{d}_N \frac{\hat{w}_N'(z)}{1+\sigma^2 c_N \hat{m}_N(z)}$. We can show that

$$\left|\frac{1}{2\pi\mathbf{i}}\oint_{\partial\mathscr{R}_y^-}\left(g_N(z)-\hat{g}_N(z)\right)\mathrm{d}z\right| \asymp 0,$$

with $\hat{w}_N(z) = z(1+\sigma^2 c_N\hat{m}_N(z))^2 - \sigma^2 c_N(1+\sigma^2 c_N\hat{m}_N(z))$.

- The previous integrand can be written as

$$g_N(z) := \mathbf{d}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z)\mathbf{I}_M\right)^{-1} \mathbf{d}_N w_N'(z)$$
$$= \mathbf{d}_N^* \mathbf{T}_N(z) \mathbf{d}_N \frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)}.$$

- From the previous result, we have the following convergence

$$m_N(z) \asymp \hat{m}_N(z) = \frac{1}{M} \operatorname{Tr} \mathbf{Q}_N(z) \quad \text{and} \quad \mathbf{d}_N^* \mathbf{T}_N(z) \mathbf{d}_N \asymp \mathbf{d}_N^* \mathbf{Q}_N(z) \mathbf{d}_N,$$

with $\mathbf{Q}_N(z) = (\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* - z\mathbf{I}_M)^{-1}$.

- Let $\hat{g}_N(z) := \mathbf{d}_N^* \mathbf{Q}_N(z) \mathbf{d}_N \frac{\hat{w}_N'(z)}{1+\sigma^2 c_N \hat{m}_N(z)}$. We can show that

$$\left| \frac{1}{2\pi \mathbf{i}} \oint_{\partial \mathscr{R}_y^-} \left(g_N(z) - \hat{g}_N(z)\right) \mathrm{d}z \right| \asymp 0,$$

with $\hat{w}_N(z) = z(1 + \sigma^2 c_N \hat{m}_N(z))^2 - \sigma^2 c_N (1 + \sigma^2 c_N \hat{m}_N(z)).$

- The new estimator is thus given by $\hat{\eta}_{new} = \frac{1}{2\pi\mathbf{i}} \oint_{\partial\mathscr{R}_y^-} \hat{g}_N(z)\mathrm{d}z$. This integral can be solved using residues theorem.
- Since 0 is separated by assumption, we deduce from the separation property that for N large enough, w.p.1

$$\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\lambda}_{M-K+1,N},\ldots,\hat{\lambda}_{M,N} \notin \mathscr{R}_y.$$

- Using argument principle, it is possible to show that for N large enough, w.p.1

$$\hat{\omega}_{1,N},\ldots,\hat{\omega}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\omega}_{M-K+1,N},\ldots,\hat{\omega}_{M,N} \notin \mathscr{R}_y,$$

  with $\hat{\omega}_{1,N} \leq \ldots \leq \hat{\omega}_{M,N}$ the solutions of the equation $1+\sigma^2 c_N \hat{m}_N(x) = 0$.
- We obtain $\hat{\eta}_N^{new} = \sum_{k=1}^M \hat{\xi}_{k,N} \mathbf{d}_N^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{d}_N$ with $(\hat{\xi}_{k,N})$ depending on $\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M,N}$ and $\hat{\omega}_{1,N},\ldots,\hat{\omega}_{M,N}$.

- The new estimator is thus given by $\hat{\eta}_{new} = \frac{1}{2\pi\mathbf{i}} \oint_{\partial\mathscr{R}_y^-} \hat{g}_N(z)\mathrm{d}z$. This integral can be solved using residues theorem.
- Since 0 is separated by assumption, we deduce from the separation property that for N large enough, w.p.1

$$\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\lambda}_{M-K+1,N},\ldots,\hat{\lambda}_{M,N} \notin \mathscr{R}_y.$$

- Using argument principle, it is possible to show that for N large enough, w.p.1

$$\hat{\omega}_{1,N},\ldots,\hat{\omega}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\omega}_{M-K+1,N},\ldots,\hat{\omega}_{M,N} \notin \mathscr{R}_y,$$

with $\hat{\omega}_{1,N} \leq \ldots \leq \hat{\omega}_{M,N}$ the solutions of the equation $1+\sigma^2 c_N \hat{m}_N(x) = 0$.
- We obtain $\hat{\eta}_N^{new} = \sum_{k=1}^M \hat{\xi}_{k,N} \mathbf{d}_N^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{d}_N$ with $(\hat{\xi}_{k,N})$ depending on $\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M,N}$ and $\hat{\omega}_{1,N},\ldots,\hat{\omega}_{M,N}$.

- The new estimator is thus given by $\hat{\eta}_{new} = \frac{1}{2\pi\mathbf{i}} \oint_{\partial\mathscr{R}_y^-} \hat{g}_N(z)\mathrm{d}z$. This integral can be solved using residues theorem.
- Since 0 is separated by assumption, we deduce from the separation property that for N large enough, w.p.1

$$\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\lambda}_{M-K+1,N},\ldots,\hat{\lambda}_{M,N} \notin \mathscr{R}_y.$$

- Using argument principle, it is possible to show that for N large enough, w.p.1

$$\hat{\omega}_{1,N},\ldots,\hat{\omega}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\omega}_{M-K+1,N},\ldots,\hat{\omega}_{M,N} \notin \mathscr{R}_y,$$

  with $\hat{\omega}_{1,N} \leq \ldots \leq \hat{\omega}_{M,N}$ the solutions of the equation $1+\sigma^2 c_N \hat{m}_N(x) = 0$.
- We obtain $\hat{\eta}_N^{new} = \sum_{k=1}^M \hat{\xi}_{k,N}\mathbf{d}_N^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{d}_N$ with $(\hat{\xi}_{k,N})$ depending on $\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M,N}$ and $\hat{\omega}_{1,N},\ldots,\hat{\omega}_{M,N}$.

- The new estimator is thus given by $\hat{\eta}_{new} = \frac{1}{2\pi\mathbf{i}} \oint_{\partial\mathscr{R}_y^-} \hat{g}_N(z)\mathrm{d}z$. This integral can be solved using residues theorem.
- Since 0 is separated by assumption, we deduce from the separation property that for N large enough, w.p.1

$$\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\lambda}_{M-K+1,N}, \ldots, \hat{\lambda}_{M,N} \notin \mathscr{R}_y.$$

- Using argument principle, it is possible to show that for N large enough, w.p.1

$$\hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M-K,N} \in \mathscr{R}_y \quad \text{and} \quad \hat{\omega}_{M-K+1,N}, \ldots, \hat{\omega}_{M,N} \notin \mathscr{R}_y,$$

  with $\hat{\omega}_{1,N} \leq \ldots \leq \hat{\omega}_{M,N}$ the solutions of the equation $1 + \sigma^2 c_N \hat{m}_N(x) = 0$.
- We obtain $\hat{\eta}_N^{new} = \sum_{k=1}^{M} \hat{\xi}_{k,N} \mathbf{d}_N^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{d}_N$ with $(\hat{\xi}_{k,N})$ depending on $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}$ and $\hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N}$.

Table of Contents

1. Introduction
2. Random matrix theory results
3. Consistent estimation of eigenspace
4. Numerical evaluations

- We evaluate the estimator in the following context:
  - $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, e^{\mathbf{i}\pi\sin(\theta)}, \ldots, e^{\mathbf{i}(M-1)\pi\sin(\theta)}]$,
  - source signals are AR(1) processes with correlation coefficient of 0.9,
  - $M = 20$, $N = 40$,
  - $K = 2$ and $\theta_1 = 16°$, $\theta_2 = 18°$.

- We evaluate the estimator in the following context:
  - $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, e^{\mathbf{i}\pi\sin(\theta)}, \ldots, e^{\mathbf{i}(M-1)\pi\sin(\theta)}]$,
  - source signals are AR(1) processes with correlation coefficient of 0.9,
  - $M = 20$, $N = 40$,
  - $K = 2$ and $\theta_1 = 16°$, $\theta_2 = 18°$.

- We evaluate the estimator in the following context:
  - $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, e^{\mathbf{i}\pi\sin(\theta)}, \ldots, e^{\mathbf{i}(M-1)\pi\sin(\theta)}]$,
  - source signals are AR(1) processes with correlation coefficient of 0.9,
  - $M = 20$, $N = 40$,
  - $K = 2$ and $\theta_1 = 16°$, $\theta_2 = 18°$.

- We evaluate the estimator in the following context:
  - $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, \mathrm{e}^{\mathbf{i}\pi\sin(\theta)}, \ldots, \mathrm{e}^{\mathbf{i}(M-1)\pi\sin(\theta)}]$,
  - source signals are AR(1) processes with correlation coefficient of 0.9,
  - $M = 20$, $N = 40$,
  - $K = 2$ and $\theta_1 = 16°$, $\theta_2 = 18°$.

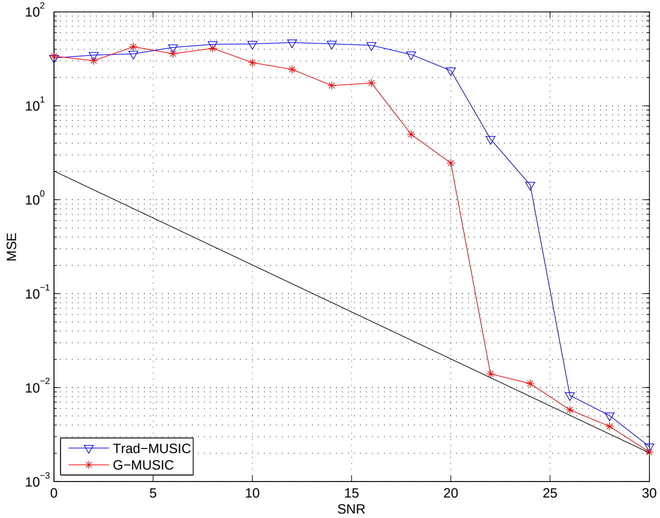


Figure: Mean of the MSE of $\hat{\theta}_1$ and $\hat{\theta}_2$ versus SNR $= 10\log(\frac{1}{\sigma^2})$

Thank you for your attention.