

On the Denoising of Cross-Spectral Matrices for (Aero)Acoustic Applications

A. Dinsenmeyer^{1,2}, J. Antoni¹, Q. Leclère¹ and A. Pereira²

¹ Laboratoire Vibrations Acoustique

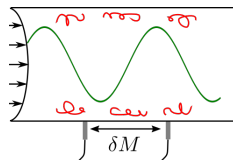
² Laboratoire de Mécanique des Fluides et d'Acoustique
Lyon, France

March 5, 2018 – 7th BeBeC



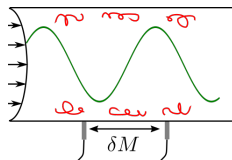
Context

- Unwanted random noise:
 - electronic, ambient, flow-induced,...
 - short correlation lengths



Context

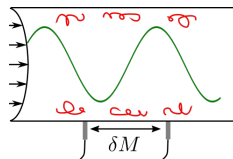
- ▶ Unwanted random noise:
 - electronic, ambient, flow-induced,...
 - short correlation lengths



- ▶ Existing denoising methods:
 - physical removal : windscreen, mic recession, porous treatment, vibrating structure filtering. . .
 - use a background noise measurement \rightarrow not always available or representative
 - wavenumber filtering \rightarrow requires high spatial sampling
 - diagonal removal \rightarrow underestimation of source level
 - exploit noise/signal properties & solve an optimization problem

Context

- ▶ Unwanted random noise:
 - electronic, ambient, flow-induced,...
 - short correlation lengths



- ▶ Existing denoising methods:
 - physical removal : windscreen, mic recession, porous treatment, vibrating structure filtering. . .
 - use a background noise measurement \rightarrow not always available or representative
 - wavenumber filtering \rightarrow requires high spatial sampling
 - diagonal removal \rightarrow underestimation of source level
 - exploit noise/signal properties & solve an optimization problem

Problem Statement

$$\underbrace{p}_{\text{measured spectra}} = \underbrace{a}_{\text{source spectra}} + \underbrace{n}_{\text{Gaussian noise}}$$

Problem Statement

$$\left\langle \underbrace{\mathbf{p}}_{\text{measured spectra}} = \underbrace{\mathbf{a}}_{\text{source spectra}} + \underbrace{\mathbf{n}}_{\text{Gaussian noise}} \right\rangle_{N_s \text{ snapshots}}$$

Cross-Spectral Matrix (covariance of Fourier component):

$$\mathbf{S}_{pp} = \frac{1}{N_s} \sum_i \mathbf{p}_i \mathbf{p}_i^H$$

- ▶ Hermitian (conjugate symmetric)
- ▶ Positive semidefinite (nonnegative eigenvalues)

Problem Statement

$$\underbrace{\langle \mathbf{p} \rangle}_{\text{measured spectra}} = \underbrace{\mathbf{a}}_{\text{source spectra}} + \underbrace{\mathbf{n}}_{\text{Gaussian noise}} \rangle_{N_s \text{ snapshots}}$$

Cross-Spectral Matrix (covariance of Fourier component):

$$\mathbf{S}_{pp} = \frac{1}{N_s} \sum_i \mathbf{p}_i \mathbf{p}_i^H$$

- ▶ Hermitian (conjugate symmetric)
- ▶ Positive semidefinite (nonnegative eigenvalues)

$$\underbrace{\mathbf{S}_{pp}}_{\text{measured CSM}} = \underbrace{\mathbf{S}_{aa}}_{\text{signal of interest}} + \underbrace{\mathbf{S}_{nn}}_{\text{unwanted noise}} + \underbrace{\mathbf{S}_{an} + \mathbf{S}_{na}}_{\text{cross-terms}}$$

- ▶ Rank of \mathbf{S}_{aa} = number of equivalent uncorrelated sources

Context – CSM properties

$$\underbrace{\mathbf{S}_{pp}}_{\text{measured CSM}} = \underbrace{\mathbf{S}_{aa}}_{\text{signal of interest}} + \underbrace{\mathbf{S}_{nn}}_{\text{unwanted noise}} + \underbrace{\mathbf{S}_{an} + \mathbf{S}_{na}}_{\text{cross-terms}}$$

For $N_s \rightarrow \infty$

Context – CSM properties

$$\underbrace{\mathbf{S}_{pp}}_{\text{measured CSM}} = \underbrace{\mathbf{S}_{aa}}_{\text{signal of interest}} + \underbrace{\cancel{\mathbf{S}_{nn}}}_{\substack{\text{unwanted noise} \\ \approx \text{diag}(\boldsymbol{\sigma}_n^2)}} + \underbrace{\mathbf{S}_{an} + \mathbf{S}_{na}}_{\text{cross-terms}}$$

For $N_s \rightarrow \infty$

- Short correlation length : off-diagonal elements of $\mathbf{S}_{nn} \rightarrow 0$

Context – CSM properties

$$\underbrace{\mathbf{S}_{pp}}_{\text{measured CSM}} = \underbrace{\mathbf{S}_{aa}}_{\text{signal of interest}} + \underbrace{\cancel{\mathbf{S}_{nn}}}_{\substack{\text{unwanted noise} \\ \approx \text{diag}(\boldsymbol{\sigma}_n^2)}} + \underbrace{\cancel{\mathbf{S}_{an}} + \cancel{\mathbf{S}_{na}}}_{\substack{\text{cross-terms} \\ \rightarrow 0}}$$

For $N_s \rightarrow \infty$

- ▶ Short correlation length : off-diagonal elements of $\mathbf{S}_{nn} \rightarrow 0$
- ▶ Independent signal/noise : cross-terms $\rightarrow 0$

How to separate signal from noise ?

- ▶ Existing methods:
 - 3 diagonal reconstruction methods
 - Robust Principal Component Analysis (RPCA)

How to separate signal from noise ?

- ▶ Existing methods:
 - 3 diagonal reconstruction methods
 - Robust Principal Component Analysis (RPCA)

- ▶ Proposed method: Probabilistic Factor Analysis

How to separate signal from noise ?

- ▶ Existing methods:
 - 3 diagonal reconstruction methods
 - Robust Principal Component Analysis (RPCA)
- ▶ Proposed method: Probabilistic Factor Analysis
- ▶ What is the influence on denoising performance of :
 - noise level,
 - number of snapshots,
 - number of sources ?

- 1 Diagonal Reconstruction
- 2 Robust Principal Component Analysis
- 3 Probabilistic Factor Analysis
- 4 Comparison

- 1 Diagonal Reconstruction
Comparison on a test case
- 2 Robust Principal Component Analysis
- 3 Probabilistic Factor Analysis
- 4 Comparison

Diagonal Reconstruction

“Remove as much noise as possible as long as denoised CSM remains non-negative”

Diagonal Reconstruction

“Remove as much noise as possible as long as denoised CSM remains non-negative”

Convex optimization (Hald, 2017)

$$\text{maximize } \|\sigma_n^2\|_1 \quad \text{subject to } \mathbf{S}_{pp} - \text{diag}(\sigma_n^2) \geq 0$$

Problem solved with CVX Matlab toolbox

Diagonal Reconstruction

“Remove as much noise as possible as long as denoised CSM remains non-negative”

Convex optimization (Hald, 2017)

$$\text{maximize } \|\sigma_n^2\|_1 \quad \text{subject to } \mathbf{S}_{pp} - \text{diag}(\sigma_n^2) \geq 0$$

Problem solved with CVX Matlab toolbox

Linear optimization (Dougherty, 2016)

$$\text{maximize } \|\sigma_n^2\|_1 \quad \text{subject to } \mathbf{V}_{(k-1)}^H \left(\mathbf{S}_{pp} - \text{diag}(\sigma_n^2)_{(k)} \right) \mathbf{V}_{(k-1)} \geq 0$$

$\mathbf{V}_{(k-1)}$: eigenvectors of $\mathbf{S}_{pp} - \text{diag}(\sigma_n^2)_{(1,\dots,k-1)}$

Solved with *linprog* Matlab function

Diagonal Reconstruction

“Remove as much noise as possible as long as denoised CSM remains non-negative”

Convex optimization (Hald, 2017)

$$\text{maximize } \|\sigma_n^2\|_1 \quad \text{subject to } \mathbf{S}_{pp} - \text{diag}(\sigma_n^2) \geq 0$$

Problem solved with CVX Matlab toolbox

Linear optimization (Dougherty, 2016)

$$\text{maximize } \|\sigma_n^2\|_1 \quad \text{subject to } \mathbf{V}_{(k-1)}^H \left(\mathbf{S}_{pp} - \text{diag}(\sigma_n^2)_{(k)} \right) \mathbf{V}_{(k-1)} \geq 0$$

$\mathbf{V}_{(k-1)}$: eigenvectors of $\mathbf{S}_{pp} - \text{diag}(\sigma_n^2)_{(1,\dots,k-1)}$

Solved with *linprog* Matlab function

Alternating Projections (Leclère et al., 2015)

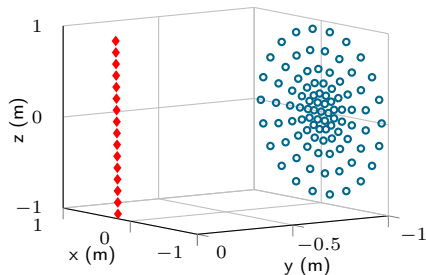
$$\mathbf{S}_{pp(k+1)} := \bar{\mathbf{S}}_{pp(0)} + \text{diag} \left(\mathbf{V}_{(k)}^H \mathbf{s}_{(k)}^+ \mathbf{V}_{(k)} \right)$$

$\mathbf{V}_{(k)}$ and $\mathbf{s}_{(k)}$: eigenvectors/values of $\mathbf{S}_{pp(k)}$

Diagonal Reconstruction – Test case

- Default parameters:
 - 20 uncorrelated free field monopoles: ♦
 - 93 receivers: ○
 - SNR: 10 dB
 - 10^4 snapshots
 - frequency: 15 kHz

- Varying parameters:
 - number of ♦ (rank of \mathbf{S}_{aa})
 - SNR
 - number of snapshots (level of extra-diagonal terms)



From a benchmark case provided by PSA3

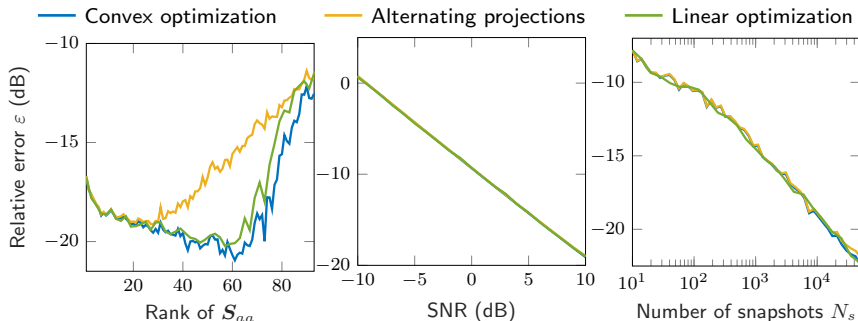
- Error on the signal CSM:

$$\varepsilon = \frac{\|\text{diag}(\mathbf{S}_{aa}) - \text{diag}(\hat{\mathbf{S}}_{aa})\|_2}{\|\text{diag}(\mathbf{S}_{aa})\|_2}$$

Diagonal Reconstruction

Default values:

20 sources	93 receivers	SNR: 10 dB	10^4 snapshots	frequency: 15 kHz
------------	--------------	------------	------------------	-------------------



Select Convex Optimization (DRec) for further comparison

- ✓ Fast, simple code
- ✓ Better performance
- ✗ Local optimization
- ✗ Denoises only auto-spectra

- 1 Diagonal Reconstruction
- 2 Robust Principal Component Analysis**
- 3 Probabilistic Factor Analysis
- 4 Comparison

RPCA

“Search \mathbf{S}_{aa} as a low rank matrix and \mathbf{S}_{nn} as a sparse matrix”

$$\text{minimize } \|\mathbf{S}_{aa}\|_* + \lambda \|\mathbf{S}_{nn}\|_1 \quad \text{subject to} \quad \mathbf{S}_{aa} + \mathbf{S}_{nn} = \mathbf{S}_{pp}$$

$\|\cdot\|_*$: nuclear norm (sum of eigenvalues: related to rank)

$\|\cdot\|_1$: ℓ_1 -norm (related to sparsity)

Solved with a proximal gradient algorithm

RPCA (Wright et al., 2009)

- ✓ Modifies the whole CSM
- ✗ Local optimization
- ✗ Choose regularization parameter:
 - L-curve criterion,
 - Generalized cross validation method,
 - Bayesian criterion, ...
- ✓ Widely used in image processing

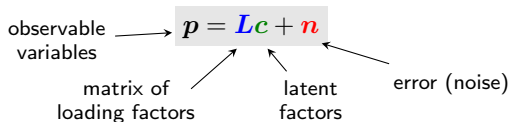
↪ For comparison :

- optimal λ (unknown on real case)
- “universal” constant parameter $\lambda = M^{-\frac{1}{2}} = 0.1$

- 1 Diagonal Reconstruction
- 2 Robust Principal Component Analysis
- 3 Probabilistic Factor Analysis**
- 4 Comparison

Probabilistic Factor Analysis

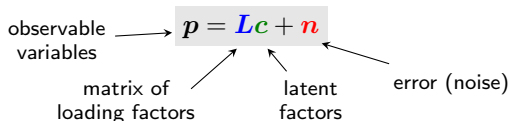
► Latent variable model



- Capture dominant correlation with fewer parameters (close to PCA)
- Extract anisotropic noise

Probabilistic Factor Analysis

► Latent variable model



- Capture dominant correlation with fewer parameters (close to PCA)
- Extract anisotropic noise

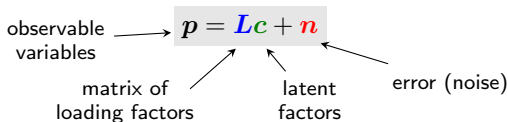
► Statistical inference: See parameters as random variables

$$L \sim \mathcal{N}_{\mathbb{C}}(0, \gamma^2) \quad c \sim \mathcal{N}_{\mathbb{C}}(0, I\alpha^2) \quad n \sim \mathcal{N}_{\mathbb{C}}(0, I\sigma^2)$$

+ non-informative priors : $\gamma^2, \alpha^2, \sigma^2 \sim \mathcal{IG}(a_{\gamma, \alpha, \sigma}, b_{\gamma, \alpha, \sigma})$

Probabilistic Factor Analysis

► Latent variable model



- Capture dominant correlation with fewer parameters (close to PCA)
- Extract anisotropic noise

► Statistical inference: See parameters as random variables

$$L \sim \mathcal{N}_{\mathbb{C}}(0, \gamma^2) \quad c \sim \mathcal{N}_{\mathbb{C}}(0, I\alpha^2) \quad n \sim \mathcal{N}_{\mathbb{C}}(0, I\sigma^2)$$

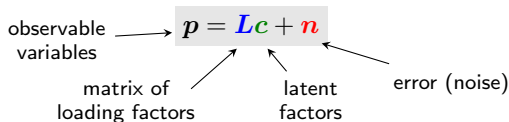
+ non-informative priors : $\gamma^2, \alpha^2, \sigma^2 \sim \mathcal{IG}(a_{\gamma, \alpha, \sigma}, b_{\gamma, \alpha, \sigma})$

► Solved using MCMC algorithm (Gibbs sampling)

Iterative draws in the marginal conditional distributions of each parameter

Probabilistic Factor Analysis

► Latent variable model



- Capture dominant correlation with fewer parameters (close to PCA)
- Extract anisotropic noise

► Statistical inference: See parameters as random variables

$$L \sim \mathcal{N}_{\mathbb{C}}(0, \gamma^2) \quad c \sim \mathcal{N}_{\mathbb{C}}(0, I\alpha^2) \quad n \sim \mathcal{N}_{\mathbb{C}}(0, I\sigma^2)$$

+ non-informative priors : $\gamma^2, \alpha^2, \sigma^2 \sim \mathcal{IG}(a_{\gamma, \alpha, \sigma}, b_{\gamma, \alpha, \sigma})$

► Solved using MCMC algorithm (Gibbs sampling)

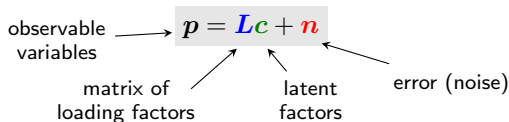
Iterative draws in the marginal conditional distributions of each parameter

► Finally, signal CSM:

$$\hat{S}_{aa} = \frac{1}{N_s} \sum_{i=1}^{N_s} L c_i c_i^H L^H$$

Probabilistic Factor Analysis

► Latent variable model



- Capture dominant correlation with fewer parameters (close to PCA)
- Extract anisotropic noise

PFA

- ✓ Global optimization
- ✓ Flexible model
- ✓ Cross-terms taken into account in the model
- ✗ Computationally expensive

- 1 Diagonal Reconstruction
- 2 Robust Principal Component Analysis
- 3 Probabilistic Factor Analysis
- 4 Comparison**

Comparison

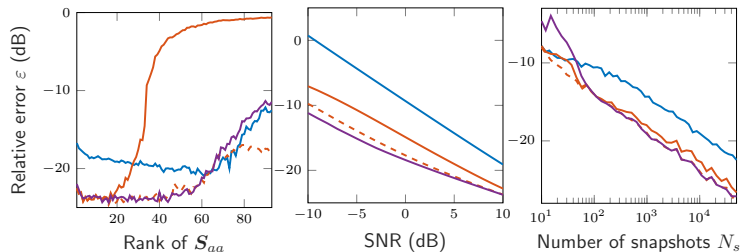
— DRec

RPCA: λ_{opt}

$\lambda = M^{-\frac{1}{2}}$

— PFA

► Homogeneous noise



Comparison

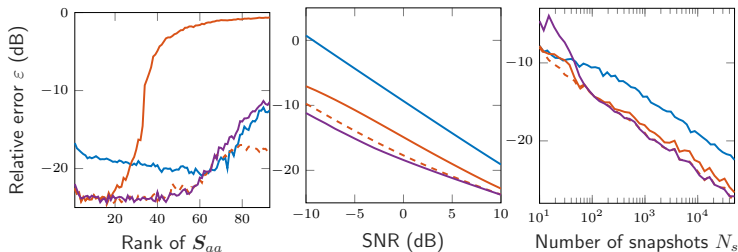
— DRec

RPCA: — λ_{opt}

— $\lambda = M^{-\frac{1}{2}}$

— PFA

► Homogeneous noise



↪ For $N_{src} \geq 0.75M$: denoising problem becomes poorly conditioned

↪ Error linearly decreases with increasing SNR

↪ Error linearly decreases with logarithmically increasing N_s

Comparison

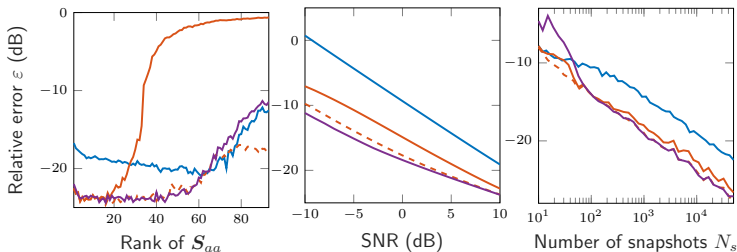
— DRec

RPCA: $\text{---} \lambda_{opt}$

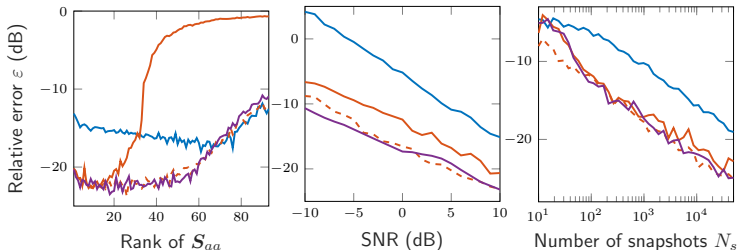
$\text{---} \lambda = M^{-\frac{1}{2}}$

— PFA

► Homogeneous noise



► Heterogeneous noise: SNR 10 dB lower on 10 random receivers



Conclusion

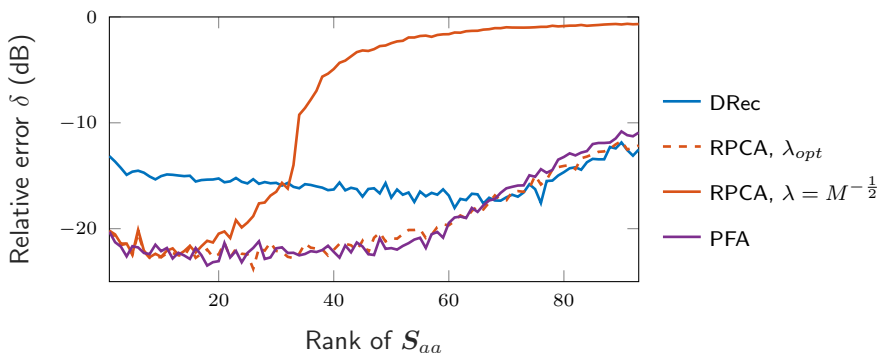
- ▶ Hard to denoise full rank CSM
- ▶ DRec: fast and simple but error 5 dB higher

- ▶ PFA
 - performance similar to RPCA using λ_{opt}
 - PFA and RPCA more robust to heterogeneous noise
 - can be solved using Expectation-Maximization algorithm
 - initialize with DRec to increase convergence speed

- ▶ Future work :
 - denoising of the whole CSM
 - adapt PFA to correlated noise
 - effect of denoising on imaging ?

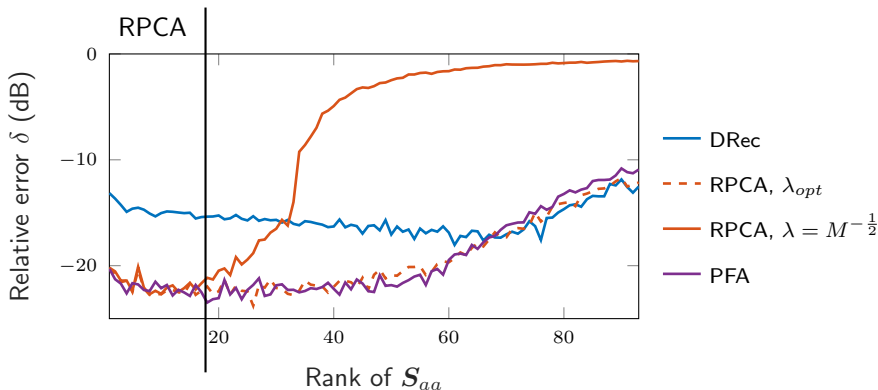
Conclusion

- Choose your denoising method according to the expected number of sources



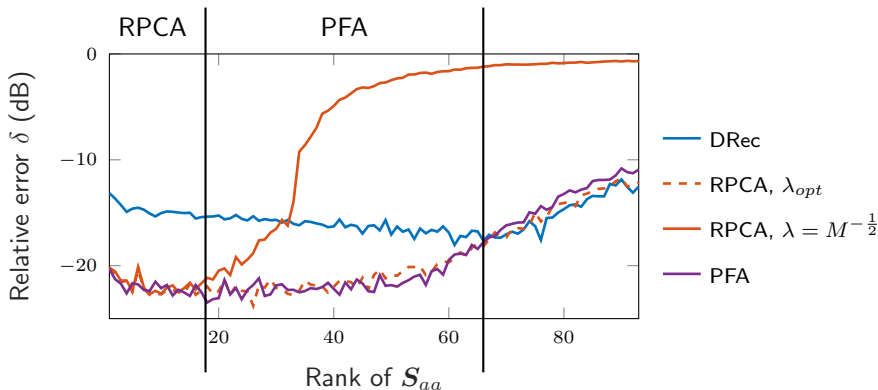
Conclusion

- Choose your denoising method according to the expected number of sources



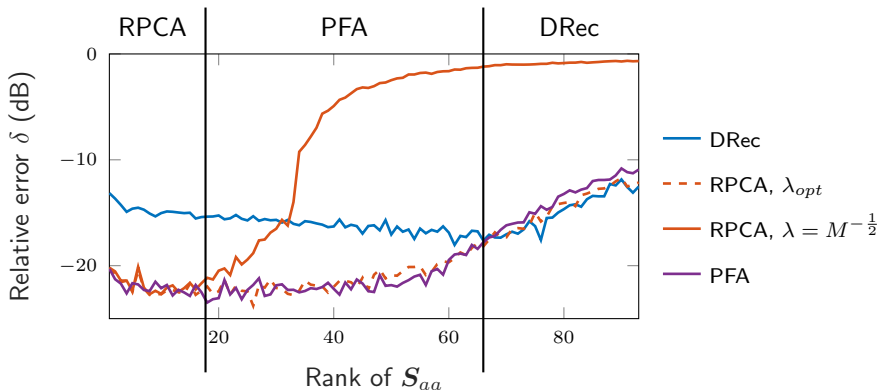
Conclusion

- Choose your denoising method according to the expected number of sources



Conclusion

- Choose your denoising method according to the expected number of sources

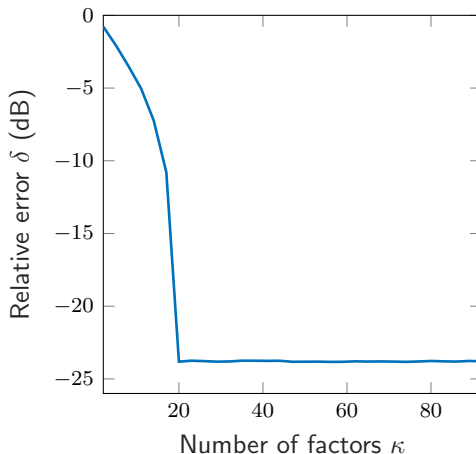


References

- R. Dougherty. Cross spectral matrix diagonal optimization. In *6th Berlin Beamforming Conference*, 02 2016.
- J. Hald. Removal of incoherent noise from an averaged cross-spectral matrix. *The Journal of the Acoustical Society of America*, 142(2): 846–854, 2017.
- Q. Leclère, N. Totaro, C. Pézerat, F. Chevillotte, and P. Souchotte. Extraction of the acoustic part of a turbulent boundary layer from wall pressure and vibration measurements. In *Novem 2015 - Noise and vibration - Emerging technologies*, Proceedings of Novem 2015, page 49046, Dubrovnik, Croatia, Apr. 2015.
- J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In *Advances in neural information processing systems*, pages 2080–2088, 2009.

PFA – Choosing the number of factor

The model enforce sparsity on factors



Bayesian inference – MAP

$$\mathbf{p} = \mathbf{L}\mathbf{c}$$

► Bayes theorem: $[\mathbf{c}|\mathbf{p}] = \frac{[\mathbf{p}|\mathbf{c}][\mathbf{c}]}{[\mathbf{p}]}$

► Maximize a posteriori density:

$$\begin{aligned}\mathbf{c} &= \arg \max_{\mathbf{c}} [\mathbf{c}|\mathbf{p}] \\ &\arg \min_{\mathbf{c}} (-\log[\mathbf{p}|\mathbf{c}] - \log[\mathbf{c}])\end{aligned}$$

- If you know which family your posterior is from
 \hookrightarrow optimization problem
- If no analytical notion of the posterior
 \hookrightarrow sampling: visit a collection of \mathbf{c} with a Markov Chain
- MCMC : performs a biased random walk to explore the distribution
 (each sample is correlated with nearby samples).

Expectation-Maximization Algorithm

Deterministic algorithm for Bayesian inference

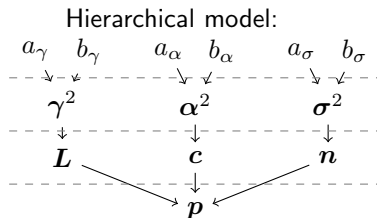
$$\mathbf{p} = \mathbf{L}\mathbf{c} + \mathbf{n}$$

1. Calculate the expected value of the likelihood function

$$Q = \mathbb{E}([\mathbf{p} \mid \mathbf{c}, \mathbf{L}, \mathbf{n}])$$

2. Find \mathbf{c} , \mathbf{L} , \mathbf{n} that maximize Q

Gibbs sampling



Gibbs sampling: update successively each variable

Require: $p, a_{\gamma}^{(0)}, b_{\gamma}^{(0)}, a_{\alpha}^{(0)}, b_{\alpha}^{(0)}, a_{\sigma}^{(0)}, b_{\sigma}^{(0)}$

for k **do**

sample c in $[c \mid p, L^{(k-1)}, \gamma^{(k-1)}, \alpha^{(k-1)}, \sigma^{(k-1)}]$

sample L in $[L \mid \text{rest}]$

sample γ^2 in $[\gamma^2 \mid \text{rest}]$

sample α^2 in $[\alpha^2 \mid \text{rest}]$

sample σ^2 in $[\sigma^2 \mid \text{rest}]$

end for