One or Two Components ? The Singular Spectrum Analysis answer

J. Harmouche[1] D. Fourer[2] P. Flandrin[1] F. Auger[2] P. Borgnat[1]

[1]Laboratoire de Physique (CNRS-ENS de Lyon), Lyon, France
{jinane.harmouche,patrick.flandrin,pierre.borgnat}@ens-lyon.fr

[2]LUNAM Université (Université de Nantes-IREENA), Saint-Nazaire, France
{dominique.fourer,francois.auger}@univ-nantes.fr

Abstract

Conventionally, Singular Spectral Analysis (SSA) is seen as a technique of decomposition of a signal into periodic components, trend and noise. As a complement to the separability question that was recently addressed by Empirical Modal Decomposition (EMD) [1] and "synchrosqueezing" [2], this work is concerned with the case where the components may be non-stationary. A characterization of the components separability is first proposed through the spectral interpretation of the singular spectrum. A new clustering solution is then offered for an automatic selection of the singular values that allows obtaining these components.

Introduction

SSA principle

- It is the analysis of a time series $(s_1, s_2, ..., s_N)$ using the spectrum of the singular values of its "trajectory matrix".
- The Hankel trajectory matrix S, with L rows and K columns ($K = N - L + 1$), is

$$S = \begin{pmatrix} s_1 & s_2 & \dots & s_K \\ s_2 & s_3 & \dots & s_{K+1} \\ \vdots & \vdots & \vdots & \vdots \\ s_L & s_{L+1} & \dots & s_{K+L-1} \end{pmatrix}.$$

where L is the embedding dimension and the sole parameter for the SSA algorithm.
- The algorithm consists of two main steps [3] :

Decomposition : the SVD of the trajectory matrix S
$\Rightarrow S = \sum_{i=1}^{r} S_i$ with $S_i = \sigma_i U_i V_i^T$.

Reconstruction : (1) diagonal averaging of $S_i \Rightarrow \mathcal{S}_i$
(2) grouping of $\mathcal{S}_i$ into meaningful components,
$\Rightarrow$ group $I_j : \mathcal{Y}_j = \sum_{i \in I_j} \mathcal{S}_i$.

Separability study

Stationary case of two harmonics

$$s_n = \cos(2\pi\lambda_0 n) + H\cos(2\pi\lambda_1 n + \phi) + \sigma\epsilon_n,$$

ϵ_n is a Gaussian white noise of variance equal to 1,
$N = 100$, $L = 50$, $\lambda_0 = 0.1$, $\lambda_1 \in]0, 0.1]$,
$H \in [0.01, 100]$, $\phi = \dfrac{\pi}{2}$ and σ allows for different values of Signal-to-Noise Ratio (SNR).

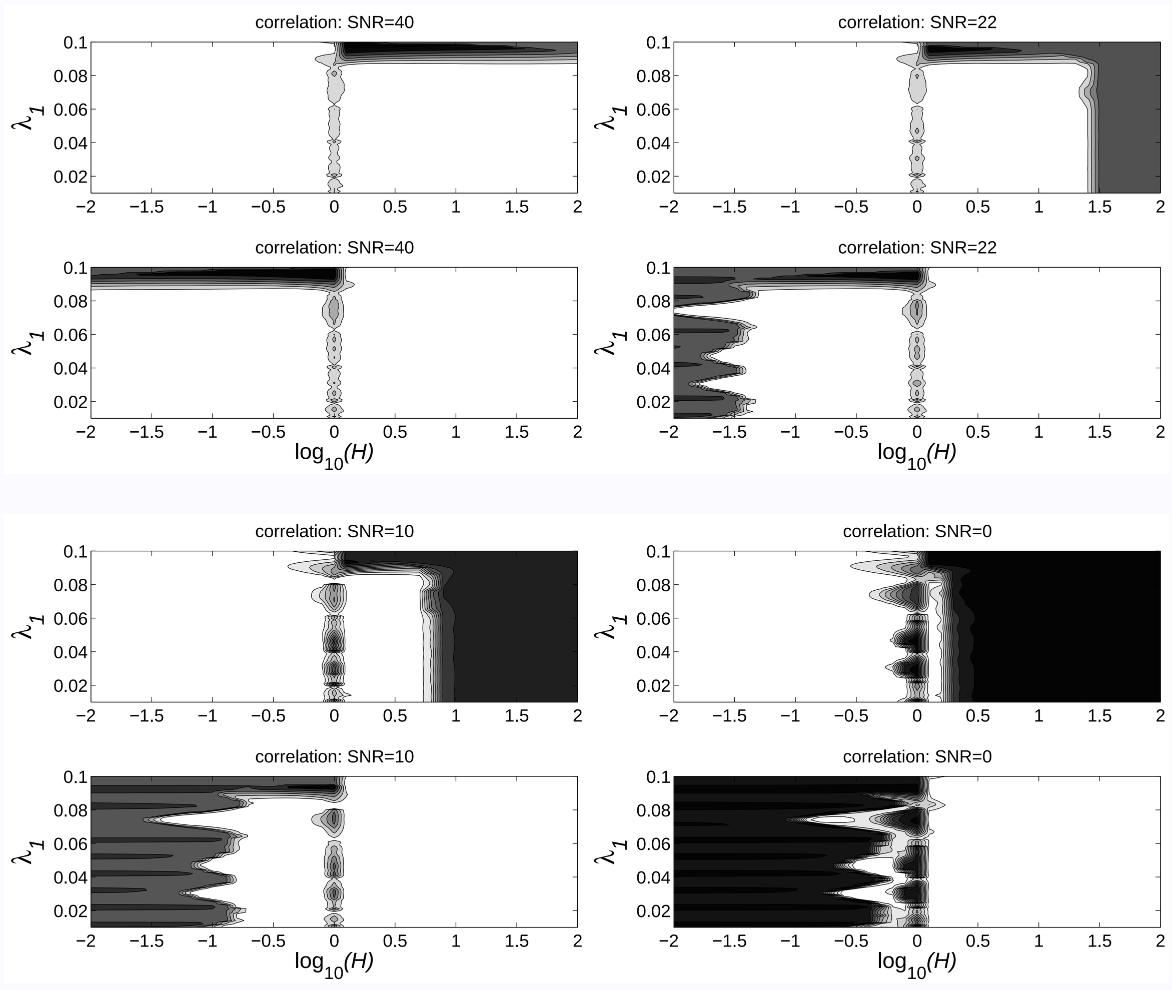

Fig.1 Stationary signal model : For each SNR, correlation coefficient between each of the harmonics (the high-frequency harmonic in the top figure) and its reconstruction. The white for total correlation and the black for total decorrelation.

The SSA can separate two sinusoids except in three situations :
- $\lambda_0 \approx \lambda_1$ gives only one pair of significant singular values
- $H = 1$ gives two pairs of equal singular values
- High noise level adds to the singular values continuous background limiting the correct identification of the sinusoids

Non-stationary case of a harmonic and a chirp

$$s_n = \cos(2\pi\lambda_0 n) + H\cos(\varphi(n)), \;\; n = 1, ..., N$$
$$\text{with } \varphi(n) = 2\pi\left(\lambda_1 n + (\delta\lambda/2N)n^2\right).$$

case 1 : $\lambda_0 < \lambda_1$
case 2 : $\lambda_1 \leq \lambda_0 \leq \lambda_1 + \delta\lambda$
case 3 : $\lambda_0 > \lambda_1 + \delta\lambda$

$L = 40$, $N = 1000$, $\lambda_1 = 0.11$ and $0 \leq \delta\lambda \leq 0.2$, the frequency λ_0 takes the values 0.1, $\lambda_1 + \delta\lambda/2$ and 0.32 in the three cases respectively.

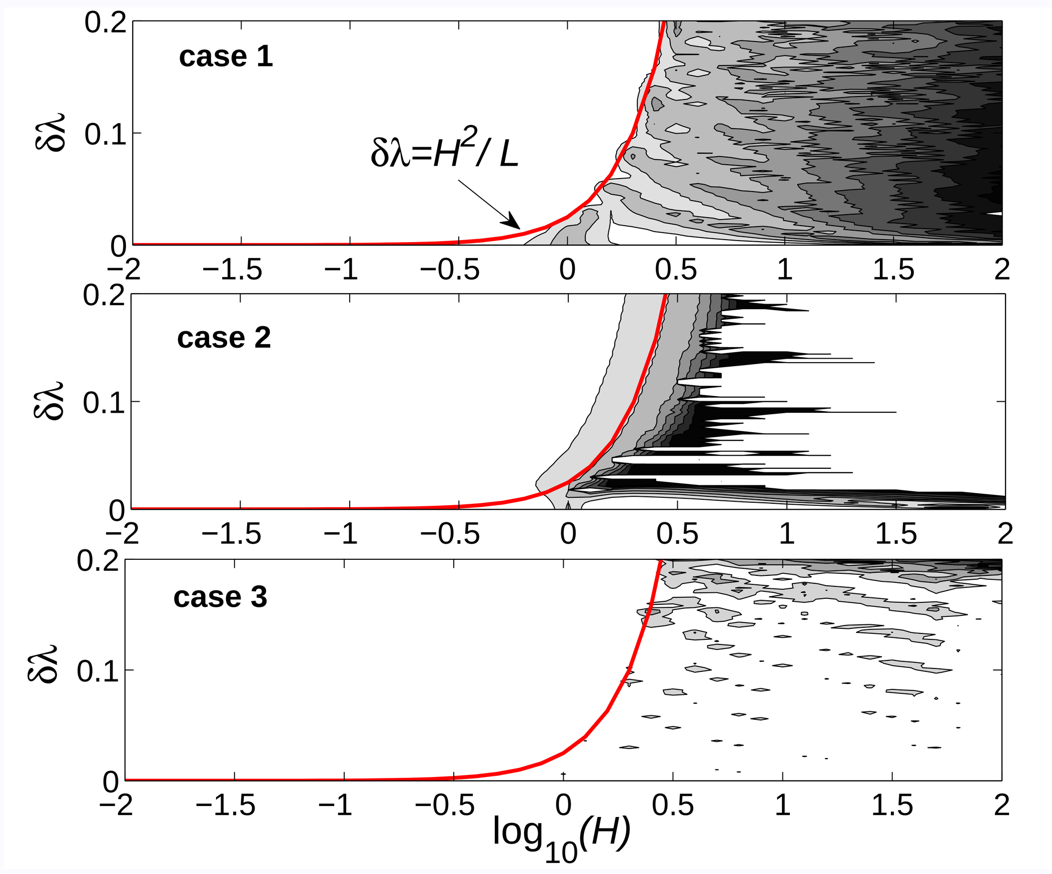

Fig.2 Non-stationary signal model : correlation coefficient between the sinusoidal component and its reconstruction obtained in supervised mode.

Results analysis

Singular spectrum interpretation

- $(s_1, ..., s_N)$: a time series
- $s \in \mathbb{R}^K$: a vector of the time series
- Fourier Transform of s : $\hat{s} = (\hat{s}_0, \hat{s}_1, ..., \hat{s}_{K-1})$

$\Rightarrow$ Under $K \gg L$, the i-th eigenvalue σ_i^2/K of the normalized covariance matrix of s can be approximated by the average value of a portion of its power spectrum, having length $l = K/L$ [4] :

$$\sigma_i^2/K \approx \frac{1}{l}\sum_{j=il-l}^{il-1}|\hat{s}_j|^2, \qquad l = \frac{K}{L}, \quad i = 1, ..., L.$$

- **A linear chirp** of normalized bandwidth $\delta\lambda \Rightarrow |\hat{s}_j|^2$ is flat and evenly distributed over $\lfloor L\delta\lambda \rfloor$ equal singular values on the interval of positive frequencies.

$\delta\lambda = H^2/L$ transition curve

Expectation : the transition curve corresponds to the situation where the level of the singular values pair associated to the pure sine wave reaches the tray of the singular values associated with the chirp.

Finding : the aforementioned spectral interpretation $\Rightarrow$ (to a factor of almost $1/l$)
- the eigenvalues associated with the sinusoid are equal to $1/2$
- the eigenvalues associated with the chirp are equal to $(H^2/2)/(L\delta\lambda)$

$$\frac{1}{2} = \frac{H^2}{2L\delta\lambda} \Rightarrow \delta\lambda = \frac{H^2}{L}$$

Unsupervised grouping approach

Hierarchical clustering

Proposal : Hierarchical clustering based classification algorithm to group signals $\mathcal{S}_i$ into meaningful components.

Algorithm description [5],[6] **:**
1. At the initialisation, $\mathcal{S}_i$ is assigned to class C_i.
2. Then, the nearest two distinct classes are merged into a new class.
3. Step 2 is repeated until the desired number of classes or the maximum dissimilarity between two elements of the class is reached.

Dissimilarity measure : $d(\mathcal{S}_i, \mathcal{S}_j)$, taking values in $[0; 1]$ and deducted from the Pearson correlation coefficient as

$$d(\mathcal{S}_i, \mathcal{S}_j) = 1 - \frac{|<\mathcal{S}_i, \mathcal{S}_j>|}{||\mathcal{S}_i|| \cdot ||\mathcal{S}_j||},$$

$$\text{with } <\mathcal{S}_i, \mathcal{S}_j> = \sum_{n=1}^{N} s_{i,n}s_{j,n} \text{ and } ||\mathcal{S}_i|| = \sqrt{<\mathcal{S}_i, \mathcal{S}_i>}.$$

Back to the separability in unsupervised mode

Supervised grouping approach : signals $\mathcal{S}_i$ are grouped by referring to the original components, assumed to be known $\Rightarrow$ **Fig.2.**

Unsupervised grouping approach : Hierarchical clustering for grouping $\Rightarrow$ **Fig.3.**

Performance : the optimal performance naturally found in the supervised mode are achieved :
- theoretical transition curve is respected
- major difference : case of spectral overlap (case 2) where the ambiguity area is transformed into an area where the separation fails everywhere

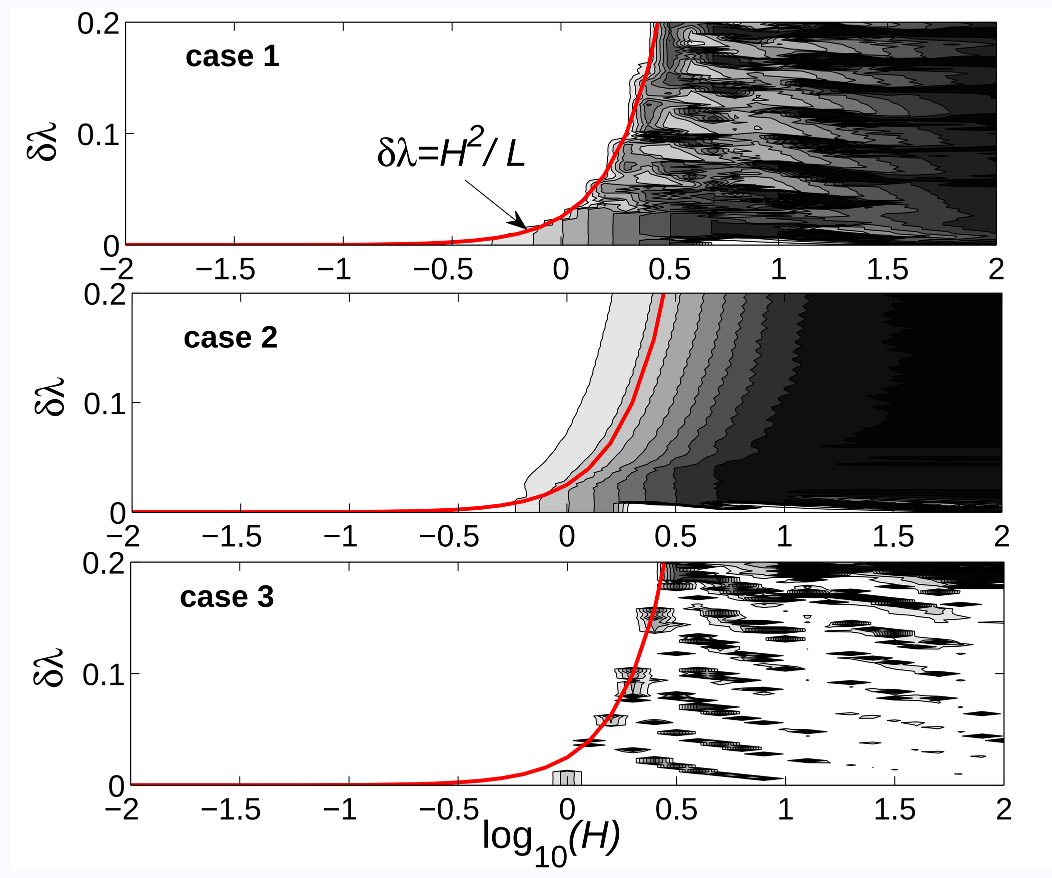

Fig. 3 Non-stationary signal model : correlation between the sinusoidal component and its reconstruction obtained with hierarchical clustering for grouping.

[1] G. Rilling and P. Flandrin, "One or two frequencies ? The Empirical Mode Decomposition answers", *IEEE Trans. Signal Process*, vol. 56, pp. 85–95, 2008.

[2] H.-T. Wu, P. Flandrin, and I. Daubechies, "One or two frequencies ? The synchrosqueezing answers", *Advances in Adaptive Data Analysis*, vol. 3, no. 1 & 2, pp. 29–39, 2011.

[3] N. Golyandina and V. Nekrutkin and A.A. Zhigljavsky, *Analysis of Time Series Structure : SSA and Related Techniques*, Chapman & Hall/CRC, 2001.

[4] E. Bozzo, R. Carniel, and D. Fasino, "Relationship between Singular Spectrum Analysis and Fourier analysis : Theory and application to the monitoring of volcanic activity", *Comp. and Math. with Applications*, vol. 60, pp. 812–820, 2010.

[5] Joe H. Ward, "Hierarchical Grouping to Optimize an Objective Function", *JASA*, vol. 58, pp. 236–244, 1963.

[6] L. Sanghoon and M.M. Crawford, "Unsupervised multistage image classification using hierarchical clustering with a bayesian similarity measure", *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 312-320, 2005.

This work was carried out with the aid of the Frensh National Research Agency as part of ASTRES project, ANR-13-BS03-0002-01.