Measurement 104 (2017) 78–88

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Structure analysis and denoising using Singular Spectrum Analysis: Application to acoustic emission signals from nuclear safety experiments

O.I. Traore [a,*], L. Pantera [b], N. Favretto-Cristini [a], P. Cristini [a], S. Viguier-Pla [d,c], P. Vieu [c]

[a] Aix-Marseille Univ, CNRS, Centrale Marseille, LMA, Marseille, France
[b] CEA, DEN, DER/SRES, Cadarache, F-13108 Saint-Paul-lez-Durance, France
[c] Institut de Mathématiques de Toulouse, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France
[d] Université de Perpignan via Domitia, LAMPS, 52 av. Paul Alduy, 66860 Perpignan Cedex 9, France

ARTICLE INFO

Article history:
Received 3 November 2016
Received in revised form 10 February 2017
Accepted 13 February 2017
Available online 4 March 2017

Keywords:
Acoustic emission
Singular Spectrum Analysis
Structural changes detection
Signal denoising
Nuclear environment

ABSTRACT

We explore the abilities of the Singular Spectrum Analysis (SSA) to characterize and denoise discrete acoustic emission signals. The method is first tested on simulated data for which different types and levels of noise are considered. It is then applied on real data recorded from nuclear safety experiments. The results show an excellent ability of the SSA to characterize the corrupted signal and to detect structural changes, even for low signal-to-noise ratio. For denoising purposes, the quality of the results depends mainly on the separability between the source signal to be estimated and the noise. However, whatever the case, the main components of the source signal are clearly identified when the components associated with the noise are removed.

1. Introduction

Reactivity Initiated Accident (RIA) is a nuclear reactor accident which involves an unexpected and very fast increase in fission rate and reactor power due to the ejection of a control rod. The power increase may damage the fuel clad and the fuel pellets of the reactor. The French Alternative Energies and Atomic Energy Commission (CEA) operates a pool-type reactor dedicated to fuel behavior study in RIA conditions.

Several non-destructive methods are used to inspect the reactor and to get information on the behavior of the fuel clad and the fuel pellets during the experiments. The acoustic emission (AE) technique is a powerful tool dedicated to structure health monitoring which has also the advantage of being simple to adapt to nuclear-oriented purposes. This technique is generally used to monitor real-time processes which emit acoustic waves. It is used to detect and/or monitor defaults and cracks in materials (e.g. [1]), and to monitor the compaction of various powders (e.g. [2–4]), including nuclear fuel powders [5,6]. Some works also focus on the correlation between acoustic signatures and physical mechanisms during tests (e.g. deformation, fragmentation, rupture...) (e.g. [7,6]). This methodology is of interest for characterizing the structure health in RIA conditions. However, since the complex environment of the nuclear reactor makes signals noisy, applying this methodology first requires an efficient signal denoising.

Several methods based on Fourier transform or time-frequency decomposition can be used for signal denoising. Most of them are based on specific assumptions on the stochastic behavior of noise. For example, in the case of the spectral subtraction, noise is supposed to be a stationary random process [8,9], whereas methods based on wavelet transform assume in general a non deterministic behavior of noise [10]. However during a nuclear safety experiment, the stochastic behavior of noise depends on the experimental protocol which may change depending on the experiment. Therefore, for such general purposes, a global denoising method which can be applied whatever the noise, and hence which makes no assumption on the stochastic behavior of noise, would be relevant.

The Singular Spectrum Analysis (SSA) method has received increasing attention since the early nineties [11,12]. Recently it has been successfully applied to various topics, for instance in geophysics [13,14] and economics [15,16]. Unlike most methods for time series analysis, SSA needs no statistical assumption on signal or noise, while performing analysis and investigation of the properties of the algorithm [17]. By using a decomposition of the signal into the sum of a small number of independent and interpretable components, SSA allows to perform various tasks such as extraction of specific components from a complex signal (noise, trend,

* Corresponding author.
E-mail addresses: toumarissiaka@gmail.com, traore@lma.cnrs-mrs.fr (O.I. Traore).

http://dx.doi.org/10.1016/j.measurement.2017.02.019
0263-2241/

seasonality ...), detection of structural changes and missing values imputation. To our best knowledge, very few works [18] have explored the ability of the SSA to analyze and denoise AE signals. This is exactly the main objective of this paper. A detailed comparison with other methods (e.g. classical pass-band filters, spectral subtraction, wavelets...) is beyond the scope of this paper, and will be the focus of a future work.

The organization of the paper is as follows. The second section reviews some basic concepts of the SSA method. Depending on the signal-to-noise ratio, the stochastic behavior of the signal and its complexity, the classical SSA may need improvements (see [17], chap.1). Furthermore, specific tasks, such as analysis of the signal structure and detection of structural changes, need to implement SSA with a more complex approach, as it is described in Section 3. This method is applied for structural analysis and denoising of simulated AE data in Section 4. The last section considers real AE signals obtained during RIA experiments.

2. Principles of the SSA method

Consider a real-valued non-zero time series $x = (x_1, \ldots, x_N)$ of length N. The main purpose of SSA is to decompose this time series into a sum of a small number of independent time series representing components of interest and residual. The SSA technique consists of two complementary stages: decomposition and reconstruction, each of them including two separate steps. In this section, we briefly summarize previous works [17,12] to which the reader may refer for much more details on the theoretical aspects.

2.1. Decomposition stage

2.1.1. Embedding

Embedding is a classical procedure in time series analysis [11,19]. It can be considered as a mapping which transfers the one-dimensional signal x into the multidimensional signals $X_1, \ldots, X_K$, where $X_i = (x_i, \ldots, x_{i+L-1})^T \in \mathbb{R}^L$ and $K = N - L + 1$. The vectors, called L-lagged vectors, are grouped into the trajectory matrix

$$\mathbf{X} = [X_1, \ldots, X_K] = \begin{pmatrix} x_1 & x_2 & \ldots & x_K \\ x_2 & x_3 & \ldots & x_{K+1} \\ \vdots & \vdots & \ldots & \vdots \\ x_L & x_{L+1} & \ldots & x_N \end{pmatrix}$$

This matrix being a Hankel matrix, all the elements along the diagonal $i + j = const$ are thus equal. The single parameter of the embedding step is the window length L which is an integer such that $2 \leqslant L \leqslant N$. The SSA technique explores the empirical distribution of the pairwise distances between the lagged vectors X_i and X_j. So the choice of the dimension L of the lagged vectors is of major importance in SSA implementation, since it determines the quality of the decomposition.

2.1.2. Singular Value Decomposition (SVD)

Based on the so-called SVD theorem [20,21], the second step of the decomposition stage consists in applying a SVD of the trajectory matrix. Denote by $\lambda_1, \ldots, \lambda_L$ the eigenvalues of the matrix $\mathbf{XX^T}$ following the decreasing order of magnitude ($\lambda_1 \geqslant \lambda_2 \geqslant \ldots \lambda_L \geqslant 0$), and by $U_1, \ldots, U_L$ the orthonormal system of the associated eigenvectors. Let $d = max\{isuchthat \lambda_i > 0\}$. If we consider $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$, the SVD of the trajectory matrix $\mathbf{X}$ can be then represented as a sum of rank-one bi-orthogonal elementary matrices $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$:

$$\mathbf{X} = \mathbf{X}_1 + \ldots + \mathbf{X}_d$$

The collection (λ_i, U_i, V_i) is called the ith eigentriple of the SVD.

2.2. Reconstruction stage

2.2.1. Grouping

The grouping procedure consists in partitioning the set of indices $\{1, \ldots, d\}$ into several groups, and in summing the matrices within each group. Considering the group $I = \{i_1, \ldots, i_p\}$ and the associated matrix $\mathbf{X}_I$ defined as $\mathbf{X}_I = \mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_p}$, the split of the set of indices $\{I_1, \ldots, I_m\}$ leads to the decomposition:

$$\mathbf{X} = \mathbf{X}_{I_1} + \ldots + \mathbf{X}_{Im}$$

The procedure for choosing the set $\{I_1, \ldots, I_m\}$ is called the eigentriple grouping.

2.2.2. Diagonal averaging

The last step of SSA transforms each matrix resulting from the grouping stage into a new signal of the same length N as the initial signal x, by using the procedure called Hankelization or diagonal averaging. It consists in transforming a given matrix $\mathbf{X}_{I_k}$ into the nearest Hankel one $\widetilde{\mathbf{X}}_{I_k}$ (see [17], chap.2):

$$\widetilde{\mathbf{X}}_{I_k} = \mathscr{H}\mathbf{X}_{I_k}$$

where $\mathscr{H}$ is the Hankel operator.

We thus obtain a signal $\tilde{x}_t^{(k)}$ corresponding to $\widetilde{\mathbf{X}}_{I_k}$ of the same length as $x = (x_1, \ldots, x_n)$, such that:

$$x_t = \sum_{k=1}^{m} \tilde{x}_t^{(k)}$$

2.3. Comments on the SSA implementation

Here, we briefly discuss the choice of the window length L which is a crucial point. We also present the concept of separability (cf. [17], chap.1 and 6). This is the central concept when studying SSA properties.

2.3.1. Choice of the window length

There is no general procedure for the selection of the window length, since the choice depends on the problem of interest and preliminary information on the signal. For example, if the signal has a periodic component with an integer period, it is convenient to consider the window length proportional to the period, in order to get a better separability of this periodic component. In many cases, the value of L should be chosen large enough, so that each L-lagged vector incorporates an essential part of the behavior of the initial signal $x = (x_1, \ldots, x_N)$. Theoretical results show that L should be large enough, but no greater than $N/2$ [17]. For a wide class of time series, the appropriate value is $median\{1, \ldots, N\}$ [22].

2.3.2. Separability

The success of the SSA decomposition strongly depends on the approximate separability of the different components of the time series. Separability between two components $x^{(1)}$ and $x^{(2)}$ of a time series characterizes how well $x^{(1)}$ and $x^{(2)}$ can be separated from each other. It can be evaluated through the weighted correlation or w-correlation:

$$\rho_{12}^{(w)} = \frac{< x^{(1)}, x^{(2)} >_w}{||x^{(1)}||_w ||x^{(2)}||_w} \quad (1)$$

Small absolute values of $\rho_{12}^{(w)}$, in particular those converging to zero, mean that $x^{(1)}$ and $x^{(2)}$ can be well separated. On the opposite, great values indicate that these components are not separable, and hence they correspond to the same component in SSA decomposition.

3. Detection of structural changes

Consider that the time series x is approximately represented as a linear combination of products of exponential, polynomial and harmonic signals, then we say that it is approximately governed by a linear recurrent formulae (LRF) (see [17], chap.3 and 5). If we assume that this property can vary with time, these variations are called structural changes or heterogeneities. For example, this is the case for AE discrete signals. The detection of these heterogeneities is of interest, especially in environment with low signal-to-noise ratio. To achieve this goal, additional tools using some steps of the classical SSA method, such as the heterogeneity matrix (H-matrix), are then needed.

The coefficients of the H-matrix are called H-indexes. The H-index $g(x^{(1)}, x^{(2)})$ characterizes the discrepancy between a base subseries $x^{(1)}$ and a test subseries $x^{(2)}$ of a time series x, of length $B > L$ and $T > L$, respectively:

$$g(x^{(1)}, x^{(2)}) = \frac{\sum_{l=1}^{T-L+1} dist^2(X_l^{(2)}, \mathscr{L}_r^{(1)})}{\sum_{l=1}^{T-L+1} ||X_l^{(2)}||^2} \tag{2}$$

Here, $\mathscr{L}_r^{(1)}$ is the linear space spanned by the set $I = \{i_1, \ldots, i_r\}$ of chosen eigenvectors of the SVD of the trajectory matrix of $x^{(1)}$. $X_1^{(2)}, \ldots, X_{T-L+1}^{(2)}$ are the L-lagged vectors of $x^{(2)}$, and $dist(X, \mathscr{L})$ is the euclidean distance between the vector $X \in \mathbb{R}^L$ and the linear space $\mathscr{L} \subset \mathbb{R}^L$.

Now, define the subseries x_{ij} of x as

$$x_{i,j} = (x_i, \ldots, x_j) \quad 1 \leqslant i < j \leqslant N.$$

The elements g_{ij} of the H-matrix $\mathbf{G} = \mathbf{G}_{B,T}$, composed of $N - B + 1$ rows and $N - T + 1$ columns, are

$$g_{ij} = g(x_{i,i+B-1}; x_{j,j+T-1})$$

The rows of the heterogeneity matrix can be described as the homogeneity of the time series x (more precisely, its test subseries $x_{j,j+T-1}$) relative to the fixed base subseries $x_{i,i+B-1}$.

Hereafter, we will use the terms time series and signal (respectively, subseries and subsignals) interchangeably.

4. Structure analysis and denoising: application to simulated AE signals

In this section, we discuss the ability of the SSA method to process AE signals. We consider a simulated source signal s (Fig. 1a) which represents a discrete AE signal of two hits (events) of different waveform characteristics (duration, energy, counts, frequency content), respectively s^1 and s^2. The first hit s^1 represents a leak event (with a low-frequency component). It has been generated by using a harmonic of frequency 5 kHz. The second one s^2 can be associated with a rupture event with a high-frequency content (Fig. 1) and has been generated by using the mathematical model proposed in [23]:

$$s^2(t) = \frac{a}{1 + e^{-b(t-\mu_1)}} e^{-c(t-\mu_2)} \sum_{i=1}^{n} A_i cos(2\pi f_i t)$$

This model is composed of two parts. The first part is a combination of sigmoid- and exponential-type functions and describes the envelop of the signal. The second one is composed of a sum of cosine functions, each with its own magnitude and frequency, and represents the frequency component involved in the rupture phenomenon.

4.1. Analysis of the source signal

4.1.1. Decomposition stage

We have considered a window length of $L \simeq N/2 = 2000$. This value also respects the choice rule for the case of harmonic components.

Analysis of the eigenvectors resulting from the SVD step shows that the eigenvectors associated with the low-frequency component are the first ones (Fig. 2). This result matches with SSA theory,

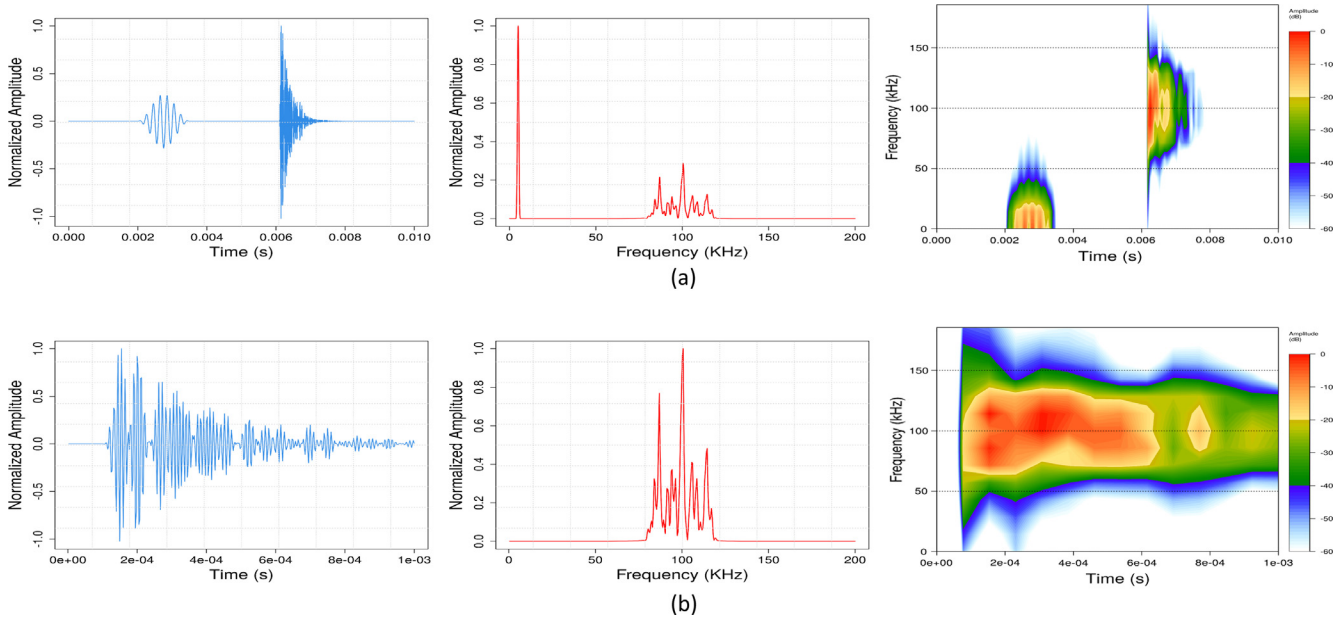

Fig. 1. Simulated source signal (a) and a focus on the second event (b). From left to right: waveform, spectrum, spectrogram.

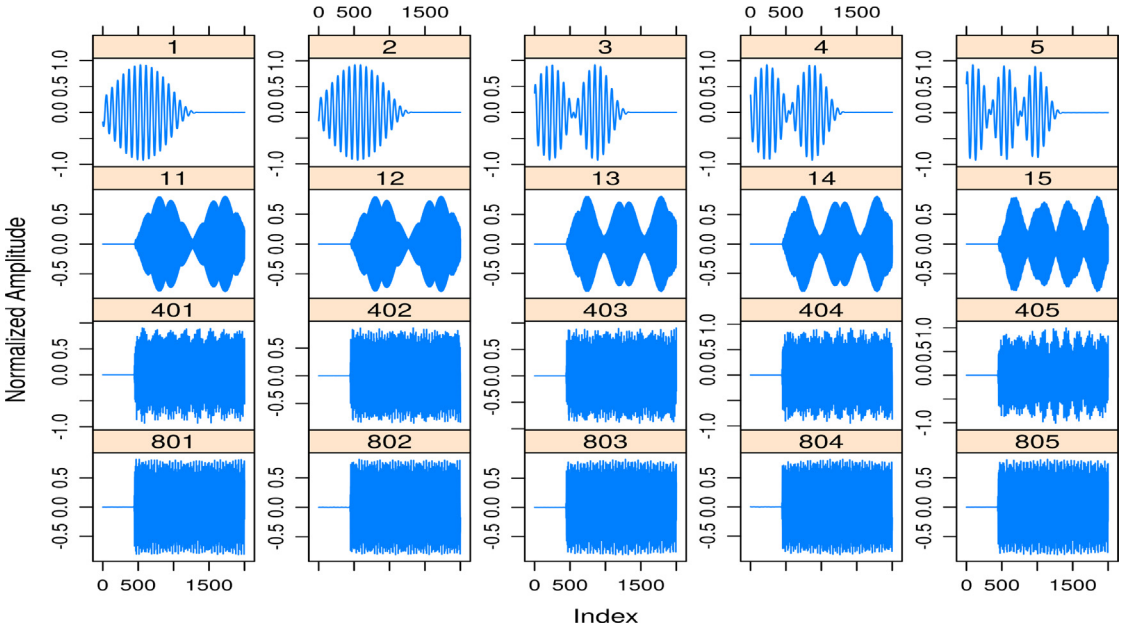

Fig. 2. Representation of some chosen eigenvectors U_i of the SVD decomposition of the trajectory matrix of the source signal s ($L = 2000$).

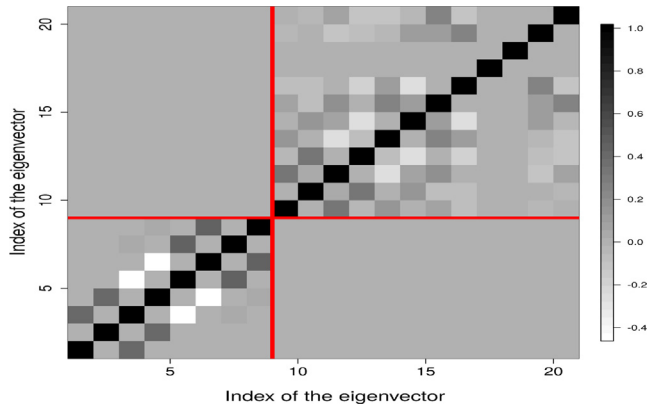

Fig. 3. Representation of the w-correlation matrix of the first twenty eigenvectors U_i of the SVD decomposition of the trajectory matrix of the source signal s ($L = 2000$).

since the position of an eigenvector is associated with its contribution to the whole signal. One can also note that some eigenvectors are highly similar by pair, indicating the presence of harmonics, which is the case for the chosen source signal here [24].

Fig. 3 represents the w-correlation matrix (see Eq. (1)) of the first twenty eigenvectors, the level of gray being proportional to the value of the w-correlation. Analysis of this matrix shows that the signal has two independent components in terms of separability. We also observe that the w-correlation between pair eigenvectors of Fig. 2 is almost equal to zero. This is explained by the fact that they do not have the same phase.

Note that additional graphical tools can also be used for analysis of the SSA decomposition results, such as the eigenvalues plot, the pairwise scatterplots, the eigenvectors spectra [17, Chapter 1].

4.1.2. Reconstruction stage

The source signal s is assumed to be noise free, so we get $\rho^{(w)}_{(s^1,s^2)} = 0$ from Eq. (1). This result confirms the excellent separability between the two events (highlighted by red lines) observed in Fig. 3. Therefore, by using the eigenvectors spectra, we can select efficiently all the eigenvectors associated with s^1 as depicted in the Fig. 4. The signal being noise free, the rest of the eigenvectors is associated with s^2. Then, we obtain the two groups of eigenvectors associated with the low- and high-frequency events, which leads to their reconstruction.

4.1.3. H-matrix analysis

The H-matrix of x highlights its low- and high-frequency events. In Fig. 5 the yellow zones associated with H-index $g_{ij} \simeq 0$ (Eq. 3) allow to detect the structural changes, and then to estimate the times of the structural changes. For a base subsignal s^1 (respectively, s^2), the heterogeneity index g_{ij} associated with the intersection with the test subsignal s^2 (respectively, s^1) is almost equal to 1, which means that there is no correlation in terms of heterogeneity between the two components.

Note that in Fig. 5 indefinite coefficients of the H-matrix have been coerced to -1 and are represented by white zones.

4.2. Analysis and denoising of noisy signals

Before any physical interpretation, the first aim in noisy environment should be to denoise the signal, i.e. extract an estimator of the source signal s from the corrupted signal x. The model associated with this problem can be written as:

$$x_t = s_t + n_t, \quad t \in \{1, \ldots, N\}$$

where x_t, s_t and n_t are the amplitudes of the corrupted signal, the source signal and the noise, respectively, at time t. In order to explore SSA abilities in this context, two types of noise n are used. Depending on the (high or low) signal-to-noise ratio and the type of noise, we highlight some important points which need to be considered for successful structural analysis and denoising. Note that in some cases, like undulations due to an external machine for instance, noise is mainly composed of one very energetic single frequency. A direct SSA and the selection of the different events by using the eigenvectors spectra then allows to perfectly reconstruct the source signal, and hence to isolate the noise. Here, the SSA can be seen as equivalent to classical pass-band filters. However, unlike

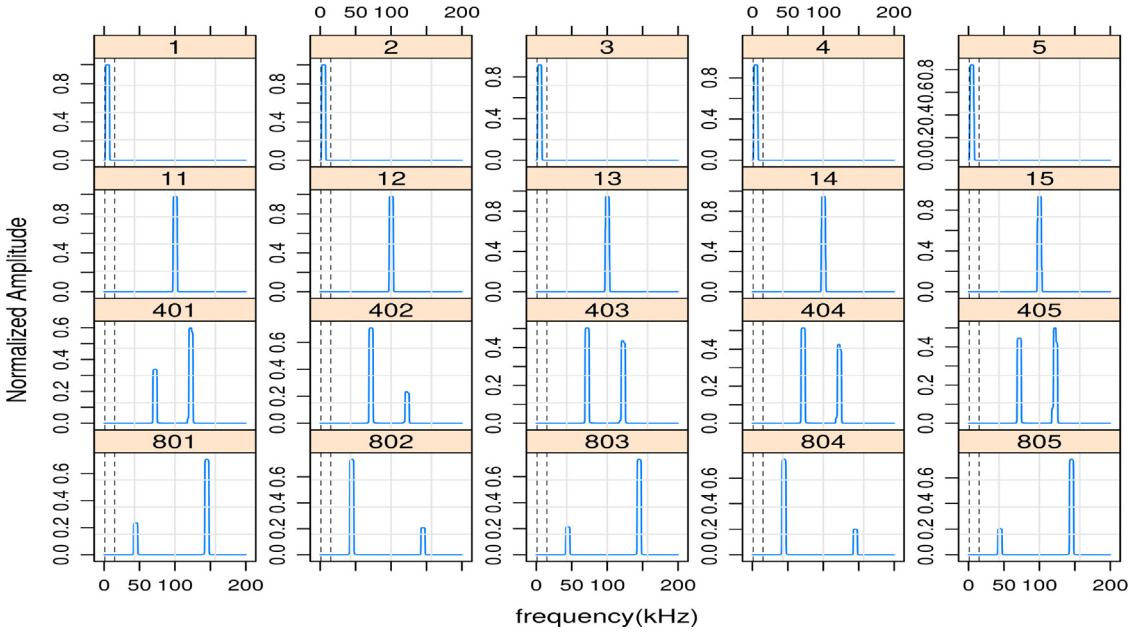

Fig. 4. Spectra of the eigenvectors U_i of the SVD decomposition of the trajectory matrix of the source signal s represented in Fig. 2.

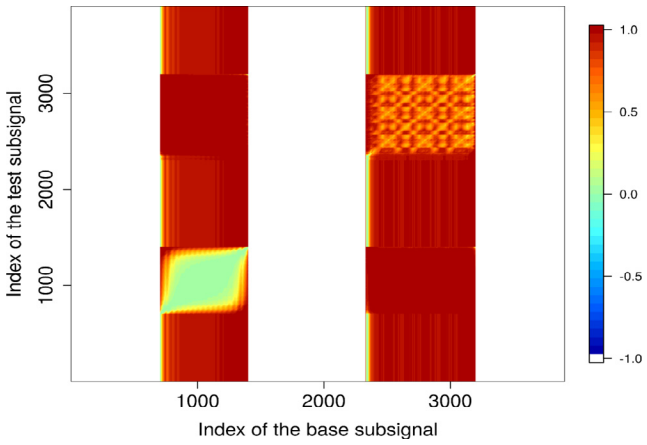

Fig. 5. H-matrix of the source signal: $B = T = 100, L = 50, \mathscr{L}_r^{(1)} = span(U_1^1, U_1^2)$.

classical filters, no cutoff frequency is taken into account. In the following, we focus on two more challenging types of noises.

The first one is a classical gaussian white noise characterized by its very wide band spectrum (Fig. 6). The second type of noise is recorded from a nuclear reactor (Fig. 7). It has a very high-frequency component (around 180 kHz) corresponding to the resonant frequency of the test device which has been used for the experiment (see Section 5). Classical wavelet-based methods are well known to be efficient in the case of random noise, such as the first one considered in our study. In contrast, they can fail in the presence of deterministic component, as for the second case of interest here. We recall that one of our goals is to experiment a global method which is able to handle any type of noise.

4.2.1. Structure analysis

Analyzing the structure of a corrupted signal first requires to pay special attention to the choice of the linear space $\mathscr{L}_r^{(1)}$ defined in Eq. (2). Indeed, depending on the type of the noise and the signal-to-noise ratio, the quality of the H-matrix is strongly affected by the choice of $\mathscr{L}_r^{(1)}$. Hereafter, we will always consider the subset of eigenvectors which span $\mathscr{L}_r^{(1)}$, namely $I = \{i_1, \ldots, i_r\} \subset \{1, \ldots, L\}$, equal to the first r components of the SVD of the base subsignals. This is equivalent to keep the r most important eigenvectors (in terms of contribution) of the base subsignals. As we show in the following, this choice can be improved depending on the obtained quality of the H-matrix.

In the case of a gaussian white noise (Fig. 6), since the noise is wide-band, it shares some frequency range with the source signal. The w-correlation between the source signal and the noise increases. Therefore, it becomes very difficult to choose eigenvectors which span $\mathscr{L}_r^{(1)}$ in such a way to fully avoid components associated with noise. Furthermore, we observe in Fig. 8a that noise

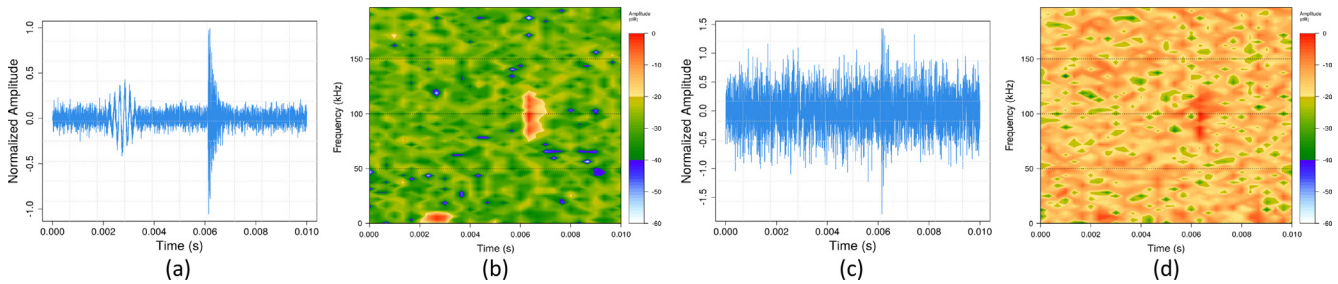

Fig. 6. Source signal corrupted by a gaussian white noise (a) and (c), and associated spectrogram (b) and (d), in high ($SNR = 15$ dB, a-b) and low ($SNR = -17$ dB, c-d) SNR context.

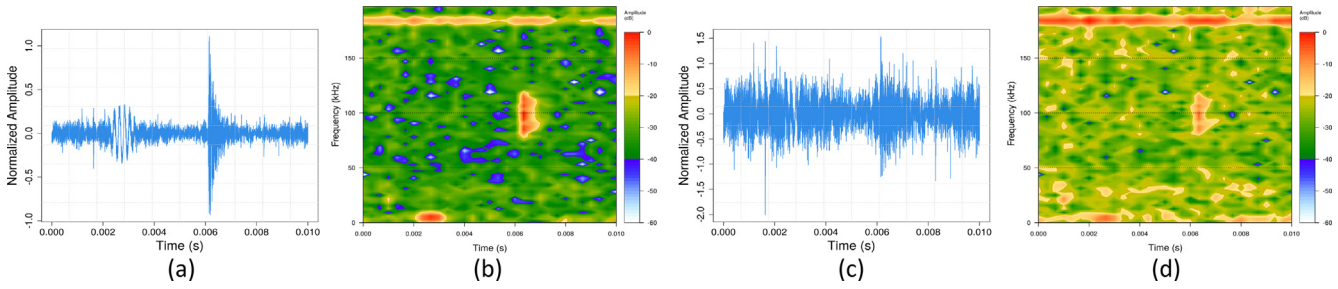

Fig. 7. Source signal corrupted by the background noise of a nuclear reactor (a) and (c), and associated spectrogram (b) and (d), in high ($SNR = 15$ dB, a-b) and low ($SNR = -17$ dB, c-d) SNR context.

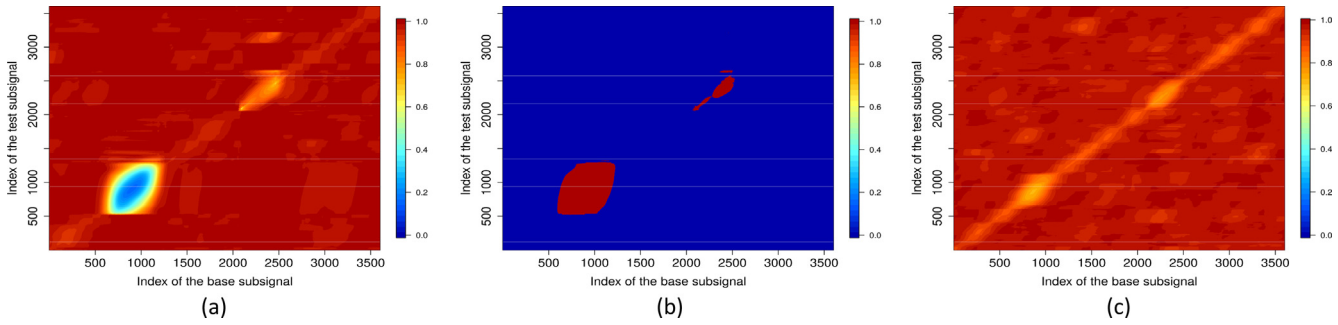

Fig. 8. H-matrix of the source signal corrupted by a gaussian white noise in high SNR context (SNR = 15 dB) before (a) and after homogenization (b), and (c) in low SNR context (SNR = −17 dB). The source signals are represented in Fig. 6. For both cases, $B = T = 400, L = 200$ and $\mathscr{L}_r^{(1)} = span(U_1^1, \ldots, U_1^6)$.

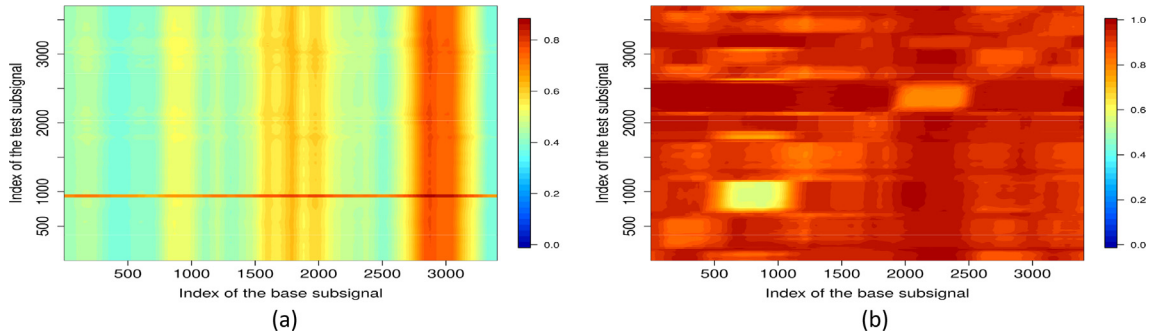

Fig. 9. Illustration of the improvement of the H-matrix quality. H-matrix (a) before and (b) after application of the SSA. Case of the source signal corrupted by a nuclear reactor noise (SNR = −17 dB). For both cases, $B = T = 350, L = 100$ and $\mathscr{L}_r^{(1)} = span(U_1^1, U_1^2)$.

contribution increases in the construction of the linear space $\mathscr{L}_r^{(1)}$ when the signal-to-noise ratio is low. This increase is emphasized by higher H-indexes for base and test subsignals corresponding to the source signal (Fig. 8c).

In some cases, it can be necessary to process the H-matrix in order to increase its quality. We propose a shrinkage which consists in replacing the H-matrix $\mathbf{G}$ by $\tilde{\mathbf{G}}$ (Fig. 8b) whose coefficients are defined by

$$\tilde{g}_{ij} = \begin{cases} 1 & \text{if } g_{ij} < threshold \\ 0 & \text{otherwise} \end{cases}$$

The choice of the threshold consists in selecting α such that:

$$threshold = mean(\mathbf{G}) \pm \alpha \sqrt{var(diag(\mathbf{G}))}$$

When the noise has its energy concentrated within a specific frequency range, as it is the case for the nuclear reactor noise examples depicted in Fig. 7, considering $\mathscr{L}_r^{(1)} = span(U_1^1, \ldots, U_1^r)$ leads to consider eigenvectors also associated with noise, and therefore to a bad quality of the H-matrix (Fig. 9a). In order to improve the H-matrix quality, we first apply a SSA to remove the main components associated with noise (Fig. 9b). These components are identified by analyzing the eigenvectors spectra and selecting those with the energy concentrated only at 180*kHz*.

4.2.2. Signal denoising by using SSA

In the case of a source signal corrupted by a gaussian white noise in a high SNR context (Fig. 6a), a deterioration of separability in comparison with a simple narrow-band noise (e.g. undulation) is observed by analyzing the *w*-correlation matrix. However, since the SNR is high and the noise very wide-band, we can suppose that the eigenvectors associated with the noise are among the last ones. Thus, by selecting the first two hundred eigenvectors in the grouping step of the SSA, we retrieve the two main components of the source signal (Fig. 10). However subsignals corresponding to pure noise segments are not equal or close to zero as they should be. This is the main consequence of the deterioration of the separability.

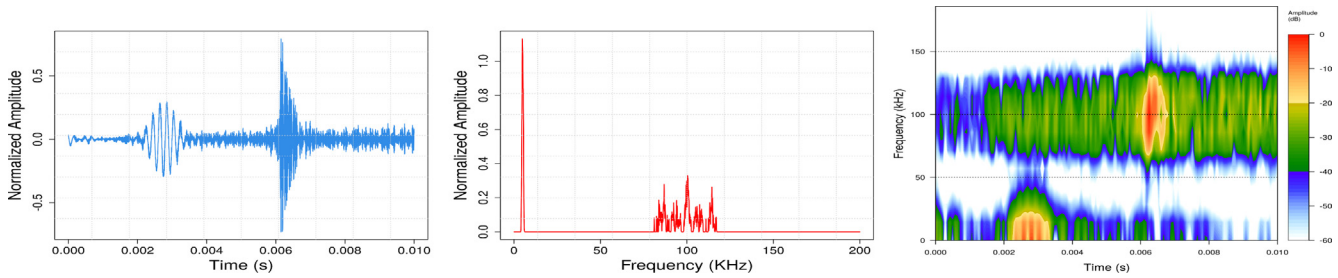

Fig. 10. Estimation of the source signal in the case of a white noise (SNR = 15 dB) by using SSA (L = 2000). From left to right: signal, spectrum and spectrogram.

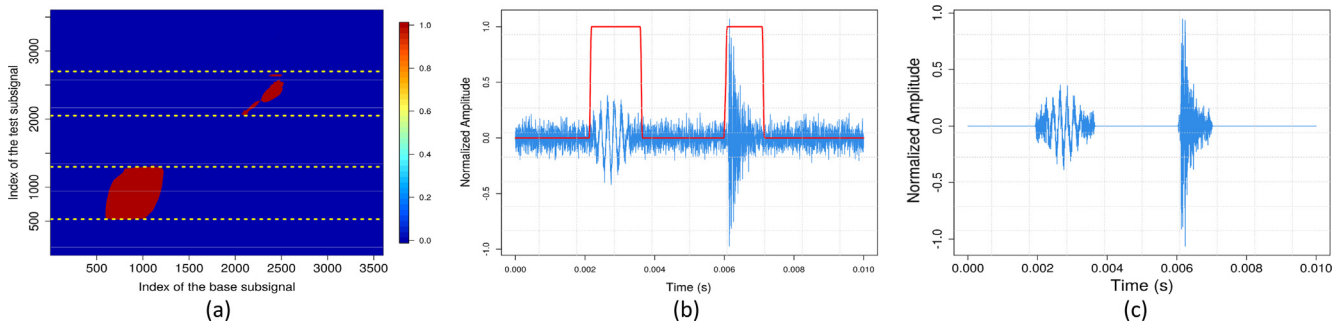

Fig. 11. First step of the denoising strategy using structural changes detection and SSA. (a) Homogenized H-matrix of the source signal, intervals associated with the test subsignals to be taken into account are indicated by yellow dash lines. (b) Detection vector (in red) applied to the noisy signal (in blue). (c) Pre-processed signal resulting from the application of the detection function to noisy signal. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

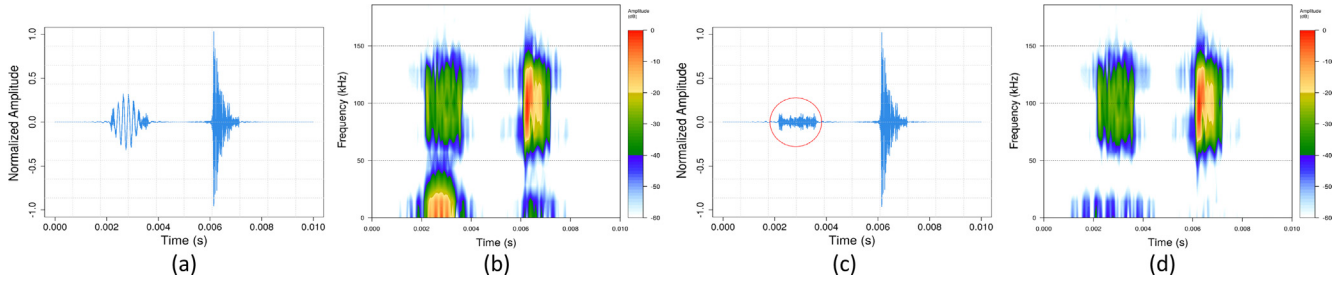

Fig. 12. Result of the classical SSA improvement by using both detection of structural changes and SSA (signals and spectrograms). (a-b) global result and (c-d) Illustration of the artifact surrounded with red in (c).

4.2.3. Signal denoising by using both detection of structural changes and SSA

In order to improve the ability of the SSA to denoise corrupted signals in the context of small w-correlations (and hence, low separability), we propose a two-steps strategy. For illustration purposes, we still consider the case of the source signal corrupted by a gaussian white noise with high SNR (Figs. 6a, b and 10).

The first step consists in estimating a function for the detection of structural changes. By fixing a given test subsignal $x_{j,j+T-1}$, with $j \in \{1,\ldots,N-T+1\}$, the function associated with the column j is the time series D^j of length $N-B+1$ defined as

$$D^j(n-B+1) = g(x_{n-B+1,n}, x_{j,j+T-1}) \quad B \leqslant n \leqslant N$$

where $g(x_{n-B+1,n}, x_{j,j+T-1})$ is the H-index between the base subsignal $x_{n-B+1,n}$ and the test subsignal $x_{j,j+T-1}$. Analyzing the H-matrix (Fig. 8b) and identifying the H-indexes associated with the source signal, we can determine the zones associated with the structural changes, and then choose the test subsignals which are relevant and need to be taken into account (Fig. 11a).

We then obtain the global function for the detection of structural changes $D = \sum_j D^j$ (Fig. 11b). Applying D to the corrupted signal leads to remove segments associated with pure noise. A pre-processed signal is thus obtained (Fig. 11c).

The second step consists in applying SSA on the pre-processed signal, in order to obtain the final estimator of the source signal. We observe that, comparatively to a simple SSA strategy (Fig. 10), the source signal estimation has been improved (Fig. 12). Analysis of the spectrogram of the estimated source signal (Fig. 12b) highlights two artifacts in comparison with the source signal (Fig. 1). These artifacts are due to an essential limitation of the SSA. Indeed, at the grouping step, choosing a given eigenvector is equivalent to selecting the frequency range associated with this eigenvector. Then, if the noise shares a part of the frequency range with a given component (for example s^2), selecting eigenvectors in order to get s^2 is equivalent to creating an artifact for s^1 (Figs. 12c and d). One solution to this problem could be to modify the classical SSA by combining it with multi-scale signal analysis methods [25].

Whatever the SNR, the proposed method always leads to the identification of the main components of the source signal. However, since the contribution of the noise increases while SNR decreases, we observe an important waveform distortion in Fig. 13. This comes from two facts. Firstly, the eigenvectors associated with the source signal become more and more difficult to isolate in the grouping step. Secondly, the impact of the artifact due to the frequency range of the source signal shared with the noise increases.

In the case of a noise associated with nuclear reactor environment, analysis of the corrupted signal spectrograms (Figs. 7b and d) shows that the frequency range associated with the noise (around 180*kHz*) has the highest energy. Therefore, the eigenvectors associated with the noise will be among the first ones after the SVD application. As for H-matrix, one can apply a first SSA to remove components mainly associated with noise, and then apply the denoising strategy presented above. Note that in this case the first SSA is equivalent to noise reduction.

5. Structure analysis and denoising: application to real AE data

During RIA experiments, the fuel rod to test is included in a test device equipped with two piezoelectric AE sensors made of lithium niobate cristal with a pass-band of 5–400 kHz and designed to work in nuclear harsh environment up to 600 degrees Celsius. They are located at the top and the bottom of the fuel rod, respectively, and are distant of 2 m (Fig. 14). A (40 dB) preamplifier is associated with each sensor, the sampling rate being 2.5 μs.

The two sensors detect the acoustic events which occur during the RIA experiments. We have no a priori assumption on the source mechanisms associated with them. After each experiment, a post-processing of the recorded signals is performed to detect the events from the whole experimental signal. One of the final goals of our work would be to identify the corresponding physical source mechanism. In this section, we explore the SSA capabilities to denoise the experimental signal and to detect structural changes (events). The identification of their exact source mechanism is beyond the scope of this paper.

Here, we focus only on the structural analysis and the denoising of the signals received at the sensor 2 during two experiments among fourteen experiments performed since 1993. Note that the methodology still remains valid for the other sensor.

The noises associated with these two experiments have different structures and levels. In the case of the first experiment the noise has two main types of components (Figs. 16a and b). The first one is a very low frequency component, probably due to a specific machine used during the experiment, whereas the second component is a very high-frequency one associated with the resonant frequency of the test device. During the second experiment, only the very high-frequency component associated with the resonant frequency of the test device has been observed (Figs. 16c and d).

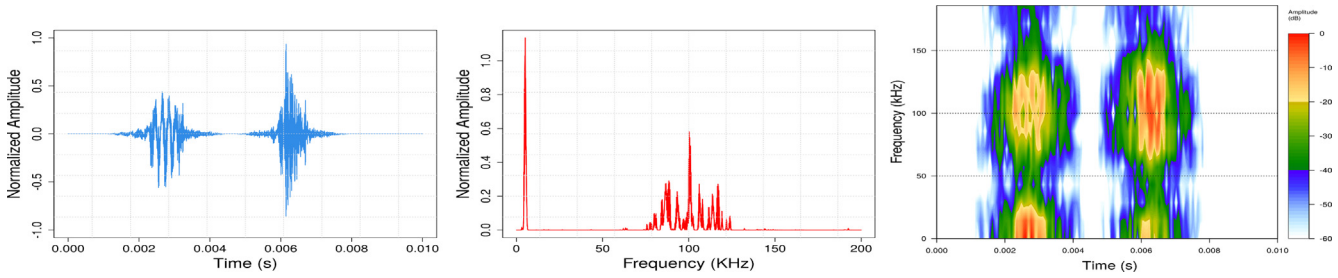

Fig. 13. Result of the denoising by using both detection of structural changes and SSA. Case of the source signal corrupted by a gaussian white noise (SNR = −17 dB). From left to right: signal, spectrum and spectrogram.

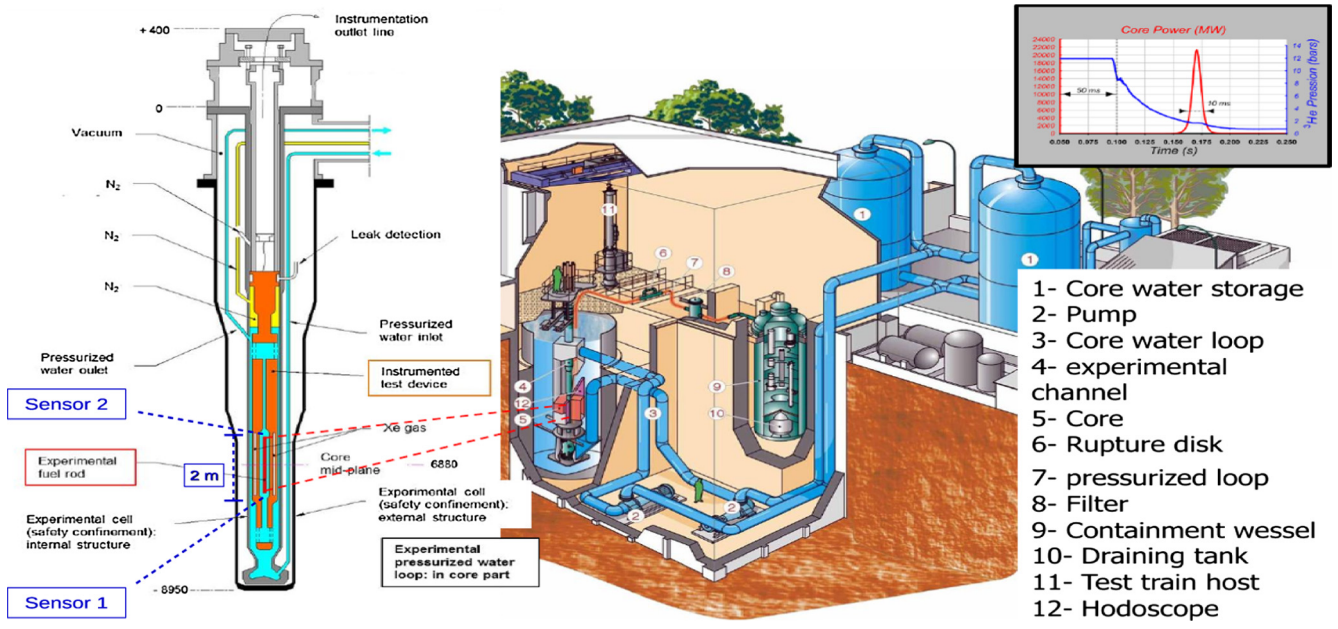

Fig. 14. Sketch of the RIA experiment reactor and the test device containing the fuel sample.

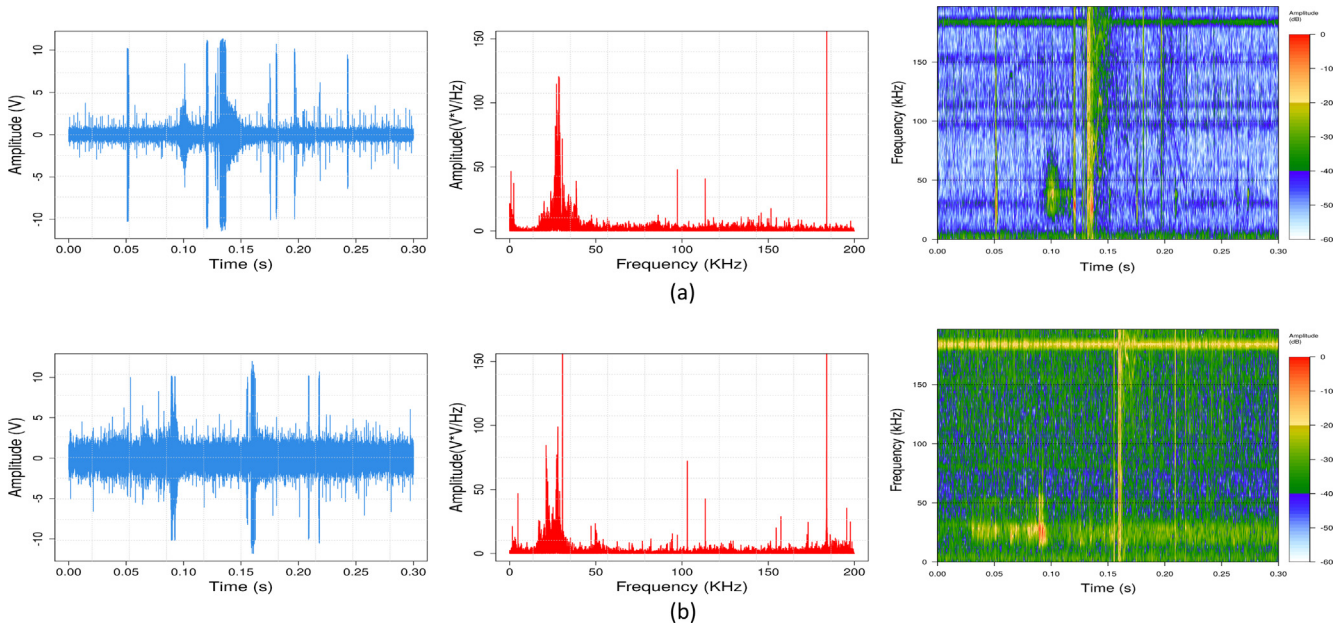

Fig. 15. From left to right: Corrupted signals recorded during the nuclear safety experiments ♯1 (a) and ♯2 (b), zoom on the associated spectrum and spectrogram.

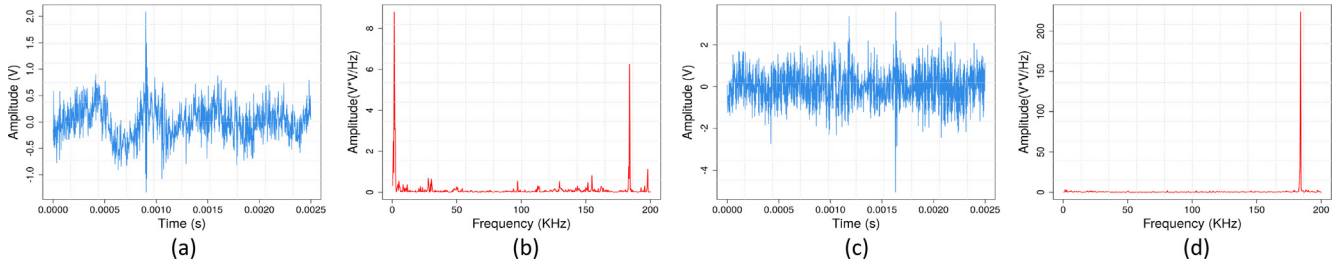

Fig. 16. Noise recorded during the two experiments of interest.

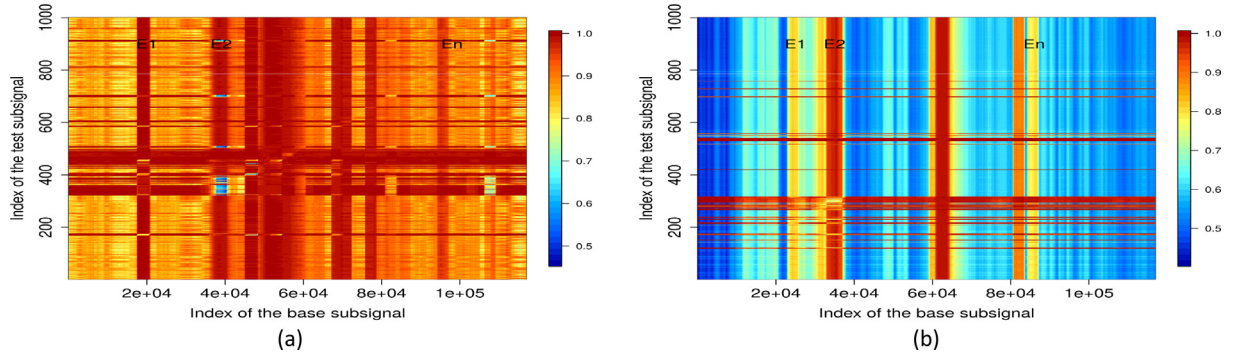

Fig. 17. H-matrices of the signals recorded from two nuclear safety experiments. (a) Experiment ♯1, (b) experiment ♯2. $T = 3000, B = 500, L = 250, \mathscr{L}_r^{(1)} = span(U_1^1, U_1^2)$.

As for the case of the simulated signals, applying SSA highlights the different structural changes in the corrupted signal (Fig. 17). In all cases, the H-matrices highlight the zones of the signal to which one has to pay attention. These zones are characterized by high H-indexes when the base subsignal coincides with a pure noise segment (zones denoted by $E_1, \ldots, E_n$ in Fig. 17). Recalling that H-indexes correspond to the discrepancies between base and test subsignals, we can conclude that in terms of heterogeneity the noise and the signal are very different.

We have used two properties of the noises in order to estimate the source signal. Firstly, the main components of noises (very low and/or high frequency) have been clearly identified by using eigenvector spectra analysis after the SVD. Secondly, since the last eigenvectors have the lowest contribution to the trajectory matrix, we can suppose that they are residual informations associated with the noises.

Analysis of the denoised signals highlights the effectiveness of the two-step denoising strategy proposed above (Fig. 18). Indeed,

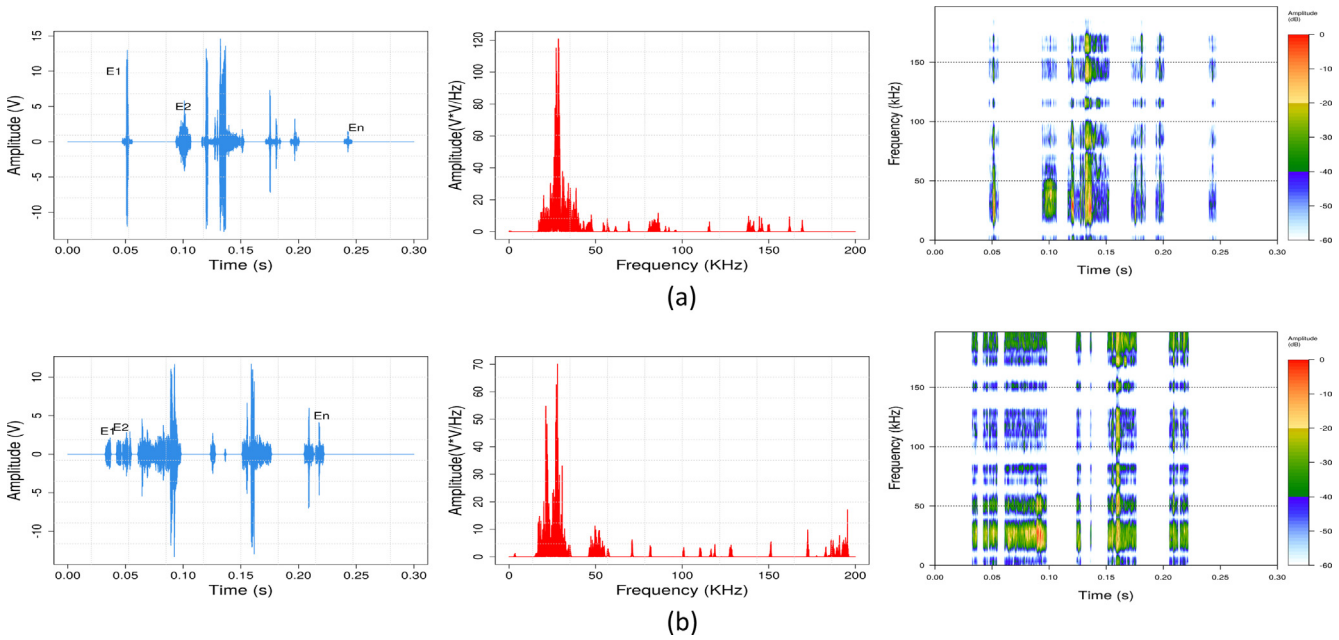

Fig. 18. From left to right: denoised signal, associated spectrum and spectrogram corresponding to the nuclear safety experiments. ♯ 1 (a) and ♯ 2 (b).

as expected, the function for the detection of structural changes allows to remove all the segments of the corrupted signal which are associated only with the noise (Figs. 15 and 18). A focus on the hits shows that there is no waveform distortion of the estimated source signal. This result is explained by the fact that the SNR is high for the experiments considered here, which is consistent with results obtained with simulated data. Furthermore, some hits (events denoted by E_n), probably associated with the source signal, have reduced energy. This is due to the fact that the energy of these events (hits) is mainly associated with noise by the SSA.

6. Conclusion

Several tools based on the classical SSA have been tested. The results obtained with real data from nuclear safety experiments match with those obtained with simulated data. In both cases, analysis of the heterogeneity matrix (H-matrix) leads to the identification of the main components of the corrupted signal, even for low signal-to-noise ratio. The H-matrix also allows the estimation of the correlation (in terms of heterogeneity) between these components. For signal denoising purposes, the SSA leads to an excellent estimation of the source signal when the separability between the noise and the source signal is such that the weighted correlation tends to zero. However, denoising becomes more difficult with increasing weighted correlation. We observe waveform distortions of the source signal and some artefacts are created when the noise and the source signal share some frequency range. In order to cope with these two problems, a denoising strategy based on two steps has been proposed. The first step, based on the H-matrix, consists in creating a function for the detection of structural changes, in order to remove all segments of the corrupted signal which are associated only with noise. In the second step, a classical SSA is applied and an eigenvector selection strategy is defined in order to isolate the last components associated with the noise. The results show that, compared to a classical application of SSA, the two-steps process significantly improves the denoising quality. However, an additional improvement is needed in order to get more accurate estimations of the source signal.

Acknowledgements

The authors thank the reviewers for constructive suggestions which have improved the paper.

References

[1] P. Antonaci, P. Bocca, D. Masera, Fatigue crack propagation monitoring by acoustic emission signal analysis, Eng. Fract. Mech. 81 (2012) 26–32.
[2] N. Parida, B. Pani, B. Kumar, Acoustic emission assisted compaction studies on iron, iron-aluminium and iron-cast iron powders, Scripta Mater. 37 (11) (1997) 1659–1663.
[3] E. Serris, L. Périer-Camby, G. Thomas, M. Desfontaines, G. Fantozzi, Acoustic emission of pharmaceutical powders during compaction, Powder Technol. 128 (2) (2002) 296–299.
[4] P. Reynaud, J. Dubois, D. Rouby, G. Fantozzi, E. Weynant, Acoustic emission monitoring of uniaxial pressing of ceramic powders, Ceram. Int. 18 (6) (1992) 391–397.
[5] L. Hégron, P. Sornay, N. Favretto-Cristini, Compaction of a bed of fragmentable particles and associated acoustic emission, IEEE Trans. Nucl. Sci. 61 (4) (2014) 2175–2181.
[6] N. Favretto-Cristini, L. Hégron, P. Sornay, Identification of the fragmentation of brittle particles during compaction process by the acoustic emission technique, Ultrasonics 67 (2016) 178–189.
[7] Y.J. Au, S. Eissa, B. Jones, Receiver operating characteristic analysis for the selection of threshold values for detection of capping in powder compression, Ultrasonics 42 (1) (2004) 149–153.
[8] S. Boll, Suppression of acoustic noise in speech using spectral subtraction, IEEE Trans. Acoust. Speech Signal Processing 27 (2) (1979) 113–120.
[9] S.V. Vaseghi, Advanced Digital Signal Processing and Noise Reduction, John Wiley & Sons, 2008.
[10] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1999.
[11] D. Broomhead, R. Jones, G. King, E. Pike, Singular system analysis with application to dynamical systems, Chaos Noise Fract. (1987) 15–27.
[12] H. Hassani, Singular spectrum analysis: methodology and comparison, J. Data Sci. 5 (2) (2007) 239–257.
[13] M. Ghil, M. Allen, M. Dettinger, K. Ide, D. Kondrashov, M. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, et al., Advanced spectral methods for climatic time series, Rev. Geophys. 40 (1) (2002).
[14] A. Serita, K. Hattori, C. Yoshino, M. Hayakawa, N. Isezaki, Principal component analysis and singular spectrum analysis of ulf geomagnetic data associated with earthquakes, Nat. Hazards Earth Syst. Sci. 5 (5) (2005) 685–689.
[15] H. Hassani, A. Zhigljavsky, Singular spectrum analysis: methodology and application to economics data, J. Syst. Sci. Complex. 22 (3) (2009) 372–394.
[16] H. Hassani, A. Soofi, A. Zhigljavsky, Predicting daily exchange rate with singular spectrum analysis, Nonlinear Anal.: Real World Appl. 11 (3) (2010) 2023–2034.
[17] N. Golyandina, V. Nekrutkin, A. Zhigljavsky, Analysis of Time Series Structure: SSA and Related Techniques, CRC Press, 2001.

[18] V. Barat, Y. Borodin, A. Kuzmin, Intelligent acoustic emission signal filtering methods, J. Acoust. Emission 28 (2010) 109–119.
[19] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, vol. 7, Cambridge University Press, 2004.
[20] V. Klema, A. Laub, The singular value decomposition: Its computation and some applications, IEEE Trans. Autom. Control 25 (2) (1980) 164–176.
[21] G. Stewart, On the early history of the singular value decomposition, SIAM Rev. 35 (4) (1993) 551–566.
[22] H. Hassani, R. Mahmoudvand, M. Zokaei, Separability and window length in singular spectrum analysis, Comp. Math. Acad. Sci. 349 (17) (2011) 987–990.
[23] D. Wotzka, Mathematical model and regression analysis of acoustic emission signals generated by partial discharges, Appl. Comput. Math. 3 (5) (2014) 225–230.
[24] R. Vautard, M. Ghil, Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, Physica D: Nonlinear Phenom. 35 (3) (1989) 395–424.
[25] P. Yiou, D. Sornette, M. Ghil, Data-adaptive wavelets and multi-scale singular-spectrum analysis, Physica D: Nonlinear Phenom. 142 (3) (2000) 254–290.