



HAL
open science

Separation of acoustical source power spectral densities with Bayesian sparsity enforcing

Daniel Blacodon, Ali Mohammad-Djafari

► **To cite this version:**

Daniel Blacodon, Ali Mohammad-Djafari. Separation of acoustical source power spectral densities with Bayesian sparsity enforcing. *Journal of Sound and Vibration*, 2020, 480, pp.115334. 10.1016/j.jsv.2020.115334 . hal-02613347

HAL Id: hal-02613347

<https://hal.science/hal-02613347>

Submitted on 20 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SEPARATION OF ACOUSTICAL SOURCE POWER SPECTRAL DENSITIES WITH BAYESIAN SPARSITY ENFORCING

Daniel Blacodon¹ and Ali Mohammad-Djafari²

¹Onera – Aerodynamics, Aeroelasticity and Acoustics Department, ONERA, Châtillon, France, Daniel.Blacodon@onera.fr

²Laboratoire des Signaux et Système, CNRS CentraleSupélec, Université Paris-Saclay, 3, Rue Joliot-Curie, 91192 Gif-sur-Yvette, France. djafari@lss.supelec.fr

ABSTRACT

Active research is ongoing to improve the design of different patterns of aircrafts including innovative devices of noise reduction often assessed during experiments conducted with scaled models in wind tunnels or in situ with real aircrafts. Source localization methods play a fundamental role to identify the source locations which are at the origin of the annoyance. Another topic for these experiments is the establishment of theoretical models, requiring a fine picture of the Power Spectral Densities (PSDs) of the main sound sources. The topic is to extract the PSDs of the primary source signals from, the PSDs of the measured mixtures. Blind signal separation techniques seem to be suited for this problem. Among the numerous existing methods, the Bayesian separation approach has the advantage of incorporating relevant information about the PSDs of the sources, and the mixing systems to help the separation process. This approach is enforced to the separation of the PSDs of primary source signals recorded by an array of microphones during tests performed in an anechoic chamber with tonal, narrow-band and broadband acoustic sources. The Signal-to-Distortion Ratio (SDR) allows to show that the separation results are better when sparsity priors are used to describe the source PSDs rather than Gaussian ones for all the scenarios of mixtures considered in the article. We demonstrate that the SRD decreases in a similar manner as the measure of the sparseness of the PSDs of the acoustic sources.

1 INTRODUCTION

The reduction of the flyover noise footprint generated by the aircraft is a priority for the civilian aeronautic industry, which in the next decade will be faced with more restrictive noise regulation in urban regions. To deal with these issues, active research by engine manufacturers is ongoing in improving the design of inlets [1], fan [2] and compressor blades [3], and nozzles [4], while airframe builders are studying landing gears [5] and flap design [6].

In order to study the efficiency of aircraft noise reduction tools the Computational Aero-Acoustics (CAA) algorithms [7] are frequently used. CAA offers a means to get an understanding of the physics of the origin in noise generation. However, aero-acoustics problems typically require a broad range of frequencies. Hence the numerical resolution to treat problems of waves with extremely short wavelengths may become a great obstacle to obtain accurate simulations with CAA.

It is likewise to apply semi-empirical models [8] which are less sophisticated than numerical methods, but induce a great advantage to provide immediate information on the acoustic behavior of acoustic sources during development of a new aircraft.

Another important step in the apprehension of the mechanisms at the origin of the noise produced by the aircrafts concerns the experiments conducted with scaled models in the anechoic wind tunnels or in situ with real aircrafts. In these practical situations, source localization methods using sensor arrays [9, 10, 11] often play a fundamental role to identify the locations of the more noisy sources. Another aim of the experiments is to permit the validation of theoretical models based on knowledge of the PSDs characteristics of the original sound sources. In this case, the problem to be solved is different from the source localization, since the issue is to extract the PSDs of the source signals from PSDs of the mixtures measured with the microphones.

Blind Signal Separation (BSS) [12] techniques seem to be suited to handle this problem. The term blind is intended to imply that such methods can separate mixed information, into unmixed ones, even if very little is known about them. Independent Component Analysis (ICA) [12] belongs to a class of BSS methods for separating data into underlying informational components, where such data can take the form of images, sounds, or telecommunication channels.

ICA assumes that the observations are instantaneous linear combinations of the source signals and that the sources are pairwise statistically independent. The mixing matrix is taken

for granted to be invertible which gives the possibility to define a separating matrix. ICA methods estimate the separation matrix which is assumed near to the inverse of the true one, up to a permutation and a scaling ambiguity. In the situations where the number of microphones can be considered greater than the number of sources ICA based techniques are quite desirable. In contrast, the Nonnegative Matrix Factorization (NMF) [13, 14] has attracted a great deal of attention in recent years when the number of sources is greater than the number of microphones and the data are nonnegative (e.g. the Spectrum Squared Magnitudes of the acoustic source signals). NMF assumes that observation data is available in matrix form \mathbf{V} which can be factorized into two unknown nonnegative matrices \mathbf{W} and \mathbf{H} (i.e. $\mathbf{V} = \mathbf{WH}$). The matrix \mathbf{W} is a dictionary of recurring patterns, assumed to be characteristic of the data, and \mathbf{H} is defined as the activation matrix aiming at approximate every pattern in \mathbf{V} with appropriate weight, onto the dictionary. One of the major fields where NMF has been used, concerns the music transcription [15, 16]. In this sort of application, the nonnegative decomposition of the spectrogram of an observed signal is done onto a dictionary of elementary spectral components representative of building sound patterns. However, NMF in its standard setting is entirely suited to single-channel data. An extension of this approach to a Multichannel NMF (MNMF) case has been developed. It owns the advantage to allow the employment of spatial information making the separation more tractable compared to NMF [17, 18, 19].

Despite success of ICA in separating with simulated signals or in applications such as electrocardiogram, image separation, it is problematic when applied to acoustic source separation because the mixing system is not simply instantaneous, but convolutive and sometimes with interferences caused by the reflections of the primary sources. ICA can be applied to separate convolutive mixtures in the frequency domain, because the convolution is equivalent to multiplication at each frequency bin. Yet, this solution involves solving an additional problem called the permutation alignment problem. Therefore, the permutations of separating matrices at each frequency should be adjusted so that the separated signal in the time domain is restored properly. Various algorithms have been proposed to solve the permutation problem [20, 21, 22]. Among them, at that place a very popular technique called Independent Vector Analysis (IVA) bears a major advantage compared to ICA since it simultaneously solves the separation and permutation problems [23, 24]. Nevertheless, it is recognized that the generative source model in IVA does not include any specific information on the spectral structures of sources, meaning that it can be mostly used for various types of

acoustic sources. In practical situations of the characterization of aircraft noise, some sources have specific spectral structures. Typically, the majority of propulsion-system noise falls into tonal (or discrete frequency) and broadband categories. It is, for example the case for the acoustic sources radiated by helicopters, which contains broadband and tonal noise components. This is also true for a supersonic jet issuing which generates a noise spectrum invariably consists of discrete and broadband components. Thus, the introduction of a suitable source model may be a solution to improve the source separation performance. This solution is explored in methods based on Independent Low-Rank Matrix Analysis (ILRMA) [25, 26] that exploit NMF decomposition to capture the spectral structures on each primary source to use them as generative model sources in IVA.

We can point out here that in practical situations, the unobserved source components that occur in aircraft noise may be not simply really complex but also partially or totally correlated. Moreover, often these components may be modelled as being a mixture of a deterministic signal “embedded” in unwanted random disturbances (i.e. background noise). Thus, it is helpful to regard all the particular characteristics of the source signals and the noise when they are known. This leads naturally to the idea of Informed PSD Separation (ISS), where the algorithm design based on Bayesian approach allows incorporating the prior knowledge in the form of the Probability Density Function (PDF) [27, 28] for the PSDs of the sources, the unwanted noise, and the mixing matrix, into the separation process. An adequate selection of prior PDFs, generally improves the separation accuracy. This is the main motivation of the choice of the Bayesian approach to resolve the problem of the separation of the PSDs of the primary sources, starting from, the PSDs of the mixtures of acoustic sources radiated by loudspeakers on an array of microphones.

In the experiment considered in this study, the acoustic sources measured during an experiment conducted in an anechoic chamber are a tone noise, a narrowband noise and a broadband noise. The scenarios examined are limited to acoustic sources pairwise mixed to verify that the separation method of their PSDs is tractable when applied to sound disturbance generated by a simplified propulsion-system composed of broadband and narrowband components.

The purpose of this article is to examine the interest to use a sparse prior distribution [29] compared to a Gaussian one to perform the separation of the power spectral densities of the primary sources. Section 2 introduces briefly the mixture model for convolutive applications. Bayesian separation approach with a particular highlight, on the Joint Maximum A Posteriori

(JMAP) used for solving out the separation problem of the PSDs is detailed in Section 3. The experiment ran out with loudspeakers to evaluate the efficiency of the Bayesian separation approach is described in Section 4. The real signals, artificially mixed and their statistical and spectral analyses are presented in Section 5. It is pointed out here, that the hybrid simulations will make it possible to control the different calculation steps implemented in the algorithms before applying the Bayesian separation approach to the PSDs of real mixtures. Section 6 describes the priors attributed to the mixing matrix, the PSDs of the acoustic signals and the likelihood that are required to solve the separation problem. JMAP solutions for the mixing scenarios considered in the article are presented in Section 7. The separation results obtained starting from, the PSDs of unreal and real mixtures are described in Section 8. The performance of PSDs separation is also assessed by computing the distortion between original PSDs and those obtained after the separation process using the Source-to-Distortion Ratio (SDR) [30] when the priors for the PSDs are all either, Gaussian or, Gaussian and Sparse. The best results are achieved when Laplace and Gaussian priors are affected to the PSDs, compared to the cases where only Gaussian priors are used. The characteristics (more or less sparsely) of the PSD affected with the sparse prior play a role in the quality of the separation measured with SDR. This effect is studied with a sparseness measure based on the ℓ_0 norm, which allows to reliably quantify the sparseness of the PSD of a source, compared to the wideness of the frequency band of interest of the PSD of another source when they are both mixed in the observations.

2 MIXTURE MODELS

The difficulty of the blind source separation task strongly depends on the way in which the signals are mixed within the physical environment. The simplest mixing scenario is termed instantaneous mixing [31], for which most early BSS algorithms were designed. However, in acoustical source separation problem we are faced to a convolutive mixing [32], because there are propagation delays between the sources and the microphones. In such applications, the degree of mixing can be very complex when propagation medium is not anechoic. Indeed, the primary sources measured by the microphones are contaminated by echoes created on reverberant surfaces which thus must be brought into account in the separation problem. However, this situation will be not dealt with this study, because the data which we will analyze are measured in an anechoic medium. The convolutive model is briefly presented in the following sections.

2.1 Convulsive model in time domain and ICA

While many algorithms have been developed for instantaneous mixing models, in many real-world applications, such as acoustic tests carried out in anechoic chamber, the mixing process is more complex. In such situation, convolutive mixing arises due to time delays resulting from sound propagation over space.

We consider the scenario where the N sources radiate acoustic waves $\mathbf{s}_v(t) = [s_{v_1}(t) \dots s_{v_N}(t)]^T$ from the locations $\mathbf{v} = [v_1 \dots v_N]$, and are impinging on M microphones placed at positions $\mathbf{r} = [r_1 \dots r_M]$ to produce the observations $x(\mathbf{r}, t)$. These latter are generally corrupted with background noise so that the output of the m^{th} microphone can be defined as:

$$x(r_m, t) = \sum_{n=1}^N \int_0^{+\infty} a_{r_m, v_n} s_{v_n}(t - \tau) d\tau + e_m(t) \quad (1)$$

where a_{r_m, v_n} are coefficients at time t of the mixing matrix $a_{\mathbf{r}, \mathbf{v}}$, characterizing the impulse responses of the propagation medium between the sources and the microphones. Hereinafter, in order to simplify the notations, $x(r_m, t)$, a_{r_m, v_n} and s_{v_n} are replaced with $x_m(t)$, $a_{m,n}$, $s_n(t)$ respectively.

We assume now, that the impulse responses can be modelled by a *FIR* (Finite Impulse Response) filter of length L and the observations sampled at the Nyquist frequency f_N . In this case, $x_m(t)$ modelled at the discrete time $t_k = \frac{k}{f_N}$ takes the form:

$$\begin{cases} x_m(k) = \sum_{n=1}^N \sum_{l=1}^L a_{m,n,l}(k) s_n(k - k_{m,n,l}) + e_m(k); & m = 1, 2, \dots, M \\ x_m(k) = \sum_{n=1}^N a_{m,n}(k) \circledast s_n(k) + e_m(k) \end{cases} \quad (2)$$

where k is the time index, $k_{m,n,l}$ denotes the propagation delay between source n and microphone m , $a_{m,n} = (a_{m,n,0}, \dots, a_{m,n,L})$ and \circledast is the convolution operator.

Eq. (2) models the acoustic mixing as a Multiple-Input Multiple-Output (MIMO) linear system (Fig.1).

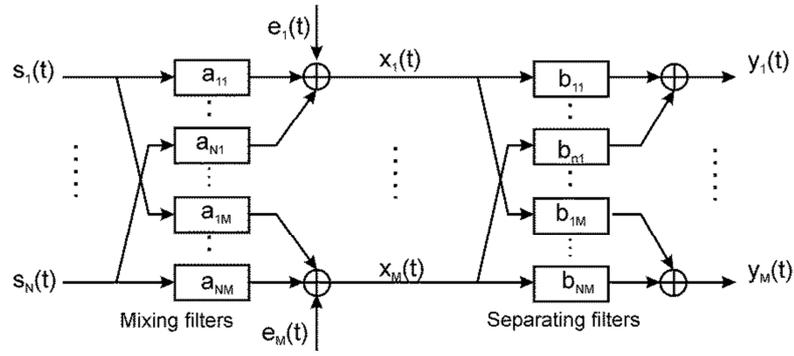


Fig.1. Block diagram of the convolutive BSS task.

The aim of multichannel blind source separation in convolutive situations is to recover the unknown primary source signals $\mathbf{s}(k)$ from the observations $\mathbf{x}(k)$, without any knowledge about the mixing filters. This task is usually achieved using a matrix of unmixing filters $b_{m,n} = (b_{m,n,0}, \dots, b_{m,n,L})$ allowing to obtain the n^{th} source signal, of the form:

$$y_n(k) = \sum_{l=0}^L \sum_{m=1}^M b_{m,n,l} x_m(k - k_{m,n,l}); \quad n = 1, \dots, N \quad (3)$$

It is possible to write $y_n(k)$ in terms of the source signals by substituting $x_m(k)$ Eq. (2) into Eq. (3), and by considering also that the separation algorithm may introduce artifacts y_{artif} , thus:

$$y_n(k) = \hat{y}_{s,n} + y_{IPS,n} + y_{BN,n} + y_{artif} \quad (4)$$

where:

$$\hat{y}_{s,n}(k) = \sum_{m=1}^M (b_{mn} \circledast a_{mn} \circledast s_n)(k) \quad (5)$$

contains the desired primary sources $s_n(k)$;

$$y_{IPS,n}(k) = \sum_{\substack{n'=1 \\ n' \neq n}}^N \sum_{m=1}^M (b_{mn'} \circledast a_{mn'} \circledast s_n)(k) \quad (6)$$

is Interfering Point Sources (*IPS*);

$$y_{BN,n}(k) = \sum_{m=1}^M (b_{mn} \circledast e_m)(k) \quad (7)$$

denotes the Background Noise (BN) components.

Eq. (4) shows that in the case of noisy convolutive mixtures, the separation filters $b_{m,n}$ must remove both the *IPS* introduced by the mixing process, the BN and the artifacts y_{artif} .

In general, BSS algorithms focus on the suppression of interfering point sources and have

only a limited capability of attenuating background noise and the artifacts. In order to help to evaluate and compare algorithms applied to blind source separation problems, it is designed in [30], three numerical performance criteria. The first one is the Source-to-Distortion Ratio (SRD):

$$SDR = 10 \log_{10} \frac{\|\hat{y}_s\|^2}{\|y_{IPS} + y_{BN} + y_{artif}\|^2} \quad (8)$$

The second one is the Source-to-Artifact Ratio (SAR):

$$SAR = 10 \log_{10} \frac{\|\hat{y}_s + y_{IPS} + y_{BN}\|^2}{\|y_{artif}\|^2} \quad (9)$$

The third one is the Source-to-Interference Ratio (SIR):

$$SIR = 10 \log_{10} \frac{\|\hat{y}\|^2}{\|y_{IPS}\|^2} \quad (10)$$

Among these three criteria, SDR measures the overall performance of the algorithms. It is chosen in the framework of this article to investigate the quality of the separation results.

2.2 Forward model for the PSDs of the observations

Time-domain algorithms can be trained to do the separation task. Nevertheless, they can be difficult to code primarily due to the multichannel convolution operations involved. There is another difficulty related to source separation algorithms based on a convolution model (Eq. (2)), which have tend to be very costly in computation time. One way to simplify the conceptualization of convolutive BSS algorithms and limit the computation costs is to solve the separation problem in the frequency domain where the linear convolution (Eq. (2)) can be written as separate multiplications for each frequency.

$$\tilde{x}_m(f) = \sum_{n=1}^N \tilde{a}_{m,n}(f) \tilde{s}_n(f) + \tilde{e}_m(f) \quad (11)$$

Where each $\tilde{x}_m(f)$ can be computed from measured $x_m(t)$. The Finite Fourier transform $\tilde{x}_m(f)$ for single records of length T of $x_m(t)$ is:

$$\tilde{x}_m(f) = \int_0^T x_m(t) e^{-2j\pi ft} dt \quad (12)$$

In a same way, $\tilde{a}_{m,n}(f)$, $\tilde{s}_n(f)$ and $\tilde{e}_m(f)$ are the Finite Fourier transform of $a_{m,n}(t)$.

The main objective of this paper is to perform the separation of the PSDs [31, 32] of the primary sources starting from the PSDs $X_m(f)$ of the observations $x_m(t)$ ($m = 1, 2, \dots, M$) defined by:

$$X_m(f) = \lim_{T \rightarrow \infty} \frac{2}{T} \mathbb{E} [|\tilde{x}_m(f)|^2] \quad (13)$$

where $\mathbb{E}[\cdot]$ is the expected value of $[\cdot]$.

Ideally, Eq.(13) involve a limit as $T \rightarrow \infty$, but the limit notation, like the frequency dependence notation, is omitted for clarity. In practice, with finite records, the limiting operation is never performed.

By substituting $\tilde{x}_m(f)$ (Eq.(11)) in Eq. (13), one obtains for $X_m(f)$:

$$X_m(f) = \frac{2}{T} \left\{ \sum_{n=1}^N \sum_{n'=1}^N \tilde{a}_{m,n} \tilde{a}_{m,n'}^* \mathbb{E}[\tilde{s}_n \tilde{s}_{n'}^*] + \sum_{n=1}^N \tilde{a}_{m,n} \mathbb{E}[\tilde{s}_n \tilde{e}_m^*] \right. \\ \left. + \sum_{n=1}^N \tilde{a}_{m,n}^* \mathbb{E}[\tilde{e}_m \tilde{s}_n^*] + \mathbb{E}[\tilde{e}_m \tilde{e}_m^*] \right\} \quad (14)$$

where * denotes the complex conjugate.

The resolution of the separation problem dealt in the article is based on several assumptions involving the primary sources and the background noise which are both supposed to be centered variables. It is considered the important special case where:

- a) $\tilde{e}_m(f)$ ($m = 1, 2, \dots, M$) is uncorrelated with each of the $s_n(t)$, ($n = 1, \dots, N$), then $\mathbb{E}[\tilde{e}_m \tilde{s}_n^*] = 0, \forall m, n$,
- b) the N acoustic sources are mutually uncorrelated, each other, then $\mathbb{E}[\tilde{s}_n \tilde{s}_{n'}^*] = S_n \delta_{n,n'}$, S_n is the PSD of $s_n(t)$ and $\delta_{n,n'}$ the Kronecker delta.

The assumptions (a) and (b) allow to reduce Eq.(13) as follows:

$$X_m = \sum_{n=1}^N A_{m,n} S_n + E_m \quad (15)$$

where $A_{m,n} = \frac{2}{T} \mathbb{E} [|\tilde{a}_{m,n}|^2]$, $S_n = \frac{2}{T} \mathbb{E} [|\tilde{s}_n|^2]$, $E_m = \frac{2}{T} \mathbb{E} [|\tilde{e}_m|^2]$

One considers here that the noise is spatially white of the same variance on all the microphones, then:

$$E_m = \sigma_E^2, (m = 1, 2, \dots, M) \quad (16)$$

Let \mathbf{X} , \mathbf{S} and \mathbf{E} are the column vectors representing the PSDs of $\mathbf{x}(t)$, $\mathbf{s}(t)$, $\mathbf{e}(t)$, and \mathbf{A} the mixing matrix:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_M \end{bmatrix}, \mathbf{A} = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{M1} & \cdots & A_{MN} \end{bmatrix}, \mathbf{S} = \begin{bmatrix} S_1 \\ \vdots \\ S_N \end{bmatrix}, \mathbf{E} = \sigma_E^2 \mathbf{I}. \quad (17)$$

From these definitions, it follows that the Forward model of the PSDs of the observations can be written in a matrix form, as:

$$\mathbf{X} = \mathbf{AS} + \mathbf{E} \quad (18)$$

2.3 Estimation of the PSDs of the observations

In practice, each element $X_m(f)$ of \mathbf{X} is obtained by computing an ensemble of estimates $|x_m^\ell(t)|^2$ from n_d continuous segments $x_m^\ell(t)$ of duration T , $((\ell - 1)T \leq t \leq \ell T, m = 1, 2, \dots, M; \ell = 1, 2, \dots, n_d)$. Each $x_m^\ell(t)$ ($m = 1, \dots, M$) is sampled at time $t_q = q\Delta t$ ($\Delta t = \frac{1}{f_N} = 0.00381 \text{ ms}$, $q = 1, \dots, Q$). In the experiment described further, the Nyquist frequency $f_N = 262144 \text{ Hz}$ and the number of samples in each data block $Q = 2048$. The frequency transformation is typically computed using a Discrete Fourier Transform (DFT) of $x_m^\ell(t_q)$, within a time frame of the prescribed length T to provide the $\tilde{x}_m^\ell(f_l)$ ($f_l = l\Delta f$, $\Delta f = 1/(Q\Delta t) = 128$, $l = 0, 1, \dots, Q - 1$). The PSDs of the $x_m(t)$ ($m = 1, 2, \dots, M$) are obtained by the ensemble averages:

$$\hat{X}_m = \frac{2}{n_d \Delta t} \sum_{\ell=1}^{n_d} |\tilde{x}_m^\ell|^2, m = 1, \dots, M. \quad (19)$$

3 BAYESIAN APPROACH FOR THE PSDs SEPARATION

The problem of PSDs separation described by Eq. (18) is particularly difficult to solve when it concerns acoustic sources. One of the ways to facilitate the separation is to take into account any information previously available on the PSDs and the mixing coefficients [33]. Bayesian methods are quite suitable, since they naturally integrate information. Moreover, in contrast from ICA methods, the Bayesian approach allows to remove unwanted noise into account in its formulation [33 - 35] to solve the separation problem from noisy observations. Thus, Bayesian methods have a remarkable aspect which their robustness to noise.

3.1 Formulation

The Bayesian PSDs separation is developed in a probabilistic framework by treating the mixing matrix, the PSDs of the sources and observations as unknown random variables. Using Bayes' rule, we have:

$$p(\mathbf{A}, \mathbf{S} | \mathbf{X}) = \frac{p(\mathbf{X} | \mathbf{A}, \mathbf{S}) p(\mathbf{A}, \mathbf{S})}{p(\mathbf{X})} \quad (20)$$

where $p(\mathbf{X} | \mathbf{A}, \mathbf{S})$, called the likelihood is in fact related to the noise \mathbf{E} in $\mathbf{X} = \mathbf{AS} + \mathbf{E}$ (Eq.(18)):

$$p(\mathbf{X} | \mathbf{A}, \mathbf{S}) = p(\mathbf{X} - \mathbf{AS}) \quad (21)$$

When the noise can be assumed to be Gaussian, then

$$p(\mathbf{X}|\mathbf{A}, \mathbf{S}) = \mathcal{N}(\mathbf{X}|\mathbf{A}\mathbf{S}, \Sigma_E) \quad (22)$$

where Σ_E is the covariance matrix of the noise \mathbf{E} which is generally assumed to be diagonal $\Sigma_E = \sigma_E^2 I$.

In that case, we can write:

$$p(\mathbf{X}|\mathbf{A}, \mathbf{S}) \propto \exp\left\{-\frac{1}{2\sigma_E^2}\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_2^2\right\} \quad (23)$$

$p(\mathbf{A}, \mathbf{S})$ is the prior on (\mathbf{A}, \mathbf{S}) which a priori can be assumed independent. It follows that:

$$p(\mathbf{A}, \mathbf{S}) = p(\mathbf{A})p(\mathbf{S}) \quad (24)$$

where $p(\mathbf{A})$ and $p(\mathbf{S})$ the prior probability distribution of \mathbf{A} and \mathbf{S} respectively.

Thus, the Bayes' rule (Eq.(20)) can be written as:

$$p(\mathbf{A}, \mathbf{S}|\mathbf{X}) = \frac{p(\mathbf{X}|\mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S})}{p(\mathbf{X})} \quad (25)$$

$p(\mathbf{X})$ in the denominator does not depend on (\mathbf{A}, \mathbf{S}) and is considered as the normalization factor allowing to write:

$$p(\mathbf{A}, \mathbf{S}|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S}) \quad (26)$$

Let us denote by $\theta = (\theta_A, \theta_S, \theta_E)$ the set of the hyperparameters associated respectively, with the variables \mathbf{A} and \mathbf{S} and \mathbf{E} of the mixing model (Eq. (18)). Thus, we have for Eq.(26):

$$p(\mathbf{A}, \mathbf{S}|\mathbf{X}, \theta) \propto p(\mathbf{X}|\mathbf{A}, \mathbf{S}, \theta_E)p(\mathbf{A}|\theta_A)p(\mathbf{S}|\theta_S) \quad (27)$$

Simple Bayesian inference for PSDs separation means to use the joint posterior $p(\mathbf{A}, \mathbf{S}|\mathbf{X}, \theta)$ to infer on both \mathbf{A} and \mathbf{S} given the data \mathbf{X} and the hyperparameters θ . A wide literature exists about Bayesian separation methods. Typical separation methods are the Maximum A posteriori (MAP) [36, 37] and the Minimum Mean Square Error (MMSE) [38].

It derives from the expression of $p(\mathbf{A}, \mathbf{S}|\mathbf{X}, \theta)$, that there are, at least two main approaches allowing the PSDs separation.

- 1) The first one is based on the Joint Maximum A Posteriori (JMAP) which allows obtaining estimates $(\hat{\mathbf{A}}, \hat{\mathbf{S}})$ of (\mathbf{A}, \mathbf{S}) using the following cost function:

$$(\hat{\mathbf{A}}, \hat{\mathbf{S}}) = \underset{(\mathbf{A}, \mathbf{S})}{\operatorname{argmax}}\{p(\mathbf{A}, \mathbf{S}|\mathbf{X}, \theta)\} \quad (28)$$

- 2) The second one consists in integrating $p(\mathbf{A}, \mathbf{S}|\mathbf{X})$ with respect to \mathbf{S} to obtain the marginal distribution $p(\mathbf{A}|\mathbf{X})$, then estimate \mathbf{A} using:

$$\hat{\mathbf{A}} = \underset{(\mathbf{A})}{\operatorname{argmax}}\{p(\mathbf{A}|\mathbf{X})\} \quad (29)$$

When \mathbf{A} is estimated, we can use it to estimate \mathbf{S} via:

$$p(\mathbf{S}|\widehat{\mathbf{A}}, \mathbf{X}) \propto p(\mathbf{X}|\widehat{\mathbf{A}}, \mathbf{S})p(\mathbf{S}) \quad (30)$$

thus,

$$\widehat{\mathbf{S}} = \underset{(\mathbf{S})}{\operatorname{argmax}}\{p(\mathbf{S}|\mathbf{X}, \widehat{\mathbf{A}})\} \quad (31)$$

In the advanced processes, the estimation of the hyperparameters $\boldsymbol{\theta}$ is also examined. However, in this work, we consider the simple case where $\boldsymbol{\theta}$ are fixed manually. Then from here, we do not mention it more.

3.2 Joint Maximum A Posteriori (JMAP) Estimation

The approach considered in this study is based on the Joint Maximum A Posterior (JMAP) to solve the separation problem of the PSDs of \mathbf{S} . The JMAP estimator Eq.(28) can be also defined as:

$$(\widehat{\mathbf{A}}, \widehat{\mathbf{S}}) = \underset{(\mathbf{A}, \mathbf{S})}{\operatorname{argmax}}\{p(\mathbf{A}, \mathbf{S}|\mathbf{X}, \boldsymbol{\theta})\} = \underset{(\mathbf{A}, \mathbf{S})}{\operatorname{argmin}}\{J(\mathbf{A}, \mathbf{S})\} \quad (32)$$

where:

$$J(\mathbf{A}, \mathbf{S}) = -\ln\{p(\mathbf{A}, \mathbf{S}|\mathbf{X}, \boldsymbol{\theta})\} = Q_1 + Q_2 + Q_3 \quad (33)$$

and

$$\begin{cases} Q_1 = -\ln(p(\mathbf{X}|\mathbf{A}, \mathbf{S}, \boldsymbol{\theta}_E)) \\ Q_2 = -\ln(p(\mathbf{A}|\boldsymbol{\theta}_A)) \\ Q_3 = -\ln(p(\mathbf{S}|\boldsymbol{\theta}_S)) \end{cases} \quad (34)$$

The first term Q_1 of the criterion (Eq. (33))

is related to the noise probability law and the forward model $\mathbf{E} = \mathbf{X} - \mathbf{AS}$. Its expression is given by $Q_1 = \frac{1}{2\sigma_E^2} \|\mathbf{X} - \mathbf{AS}\|_2^2$. The two other terms Q_2 and Q_3 depend on the priors $p(\mathbf{A})$ and $p(\mathbf{S})$. One of the easiest means to obtain the solutions by using the cost function $J(\mathbf{A}, \mathbf{S})$ is based on an alternate optimization with respect to \mathbf{A} and then to \mathbf{S} [39], thus:

$$\begin{cases} \widehat{\mathbf{A}} = \underset{\mathbf{A}}{\operatorname{argmin}}\{J(\mathbf{A}, \widehat{\mathbf{S}})\} \\ \widehat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}}\{J(\widehat{\mathbf{A}}, \mathbf{S})\} \end{cases} \quad (35)$$

Studying the convergence properties of such algorithms, in general, is not easy. There is no guaranty that they converge toward the global JMAP solutions, but may converge toward any local minimum. However, acceptable results can be obtained by choosing appropriate priors $p(\mathbf{A})$ and $p(\mathbf{S})$ or by imposing constraints on the mixing matrix and on the PSDs of the sources [27].

4 EXPERIMENT WITH LOUDSPEAKERS

In the framework of this study, an experiment has been carried out in an anechoic chamber (Fig.2) to assess the ability of the Bayesian approach to separate successfully the PSDs of the acoustic sources that may be involved in practical situations encountered in the characterization of aircraft noise.

More precisely, we are concerned to get a solution to a major problem of interest. Starting from \mathbf{X} the PSDs of the observations composed of a tonal source (i.e. \mathbf{S}_1), a narrowband source (i.e. \mathbf{S}_2) occurring in propulsion devices of aircrafts, and an unwanted broadband noise (i.e. \mathbf{S}_3) We want to truly separate out the PSDs of \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 . These latter are three loudspeakers \mathbf{S}_1 , \mathbf{S}_2 , and \mathbf{S}_3 radiating a sine wave at 4 kHz, an acoustic wave in the frequency range [3 kHz, 5 kHz] and another on in the interval [2 kHz, 10 kHz] respectively. The acoustic measurements are performed with a linear array composed of fifteen equally-spaced microphones. Moreover, three reference microphones called *Ref1*, *Ref2* and *Ref3* are also installed near the output of the loudspeaker. The acoustic waves $s_1(t)$, $s_2(t)$ and $s_3(t)$ emitted respectively by \mathbf{S}_1 , \mathbf{S}_2 , and \mathbf{S}_3 are individually measured with both the array and reference microphones. According to the previous discussions on the noise generated by propulsion devices, we only consider the three scenarios, when the pairs of loudspeakers (\mathbf{S}_1 , \mathbf{S}_2), (\mathbf{S}_1 , \mathbf{S}_3) and (\mathbf{S}_2 , \mathbf{S}_3) are active. In the present situation, we are in a reverberation-free environment with propagation delays so that the mixing model (Eq. (2)) of the three scenarios can be simplified:

$$x_m^{l,n}(t) = a_{m,l}(t)s_l(t - \tau_{m,l}) + a_{m,n}(t)s_n(t - \tau_{m,n}) + e_m(t); \quad m = 1, 2, \dots, M \quad (36)$$

where

$$\begin{cases} l, n = 1, 2 \text{ when } S_1 \text{ and } S_2 \text{ are active} \\ l, n = 1, 3 \text{ when } S_1 \text{ and } S_3 \text{ are active} \\ l, n = 2, 3 \text{ when } S_2 \text{ and } S_3 \text{ are active} \end{cases} \quad (37)$$

$\tau_{m,l}$ and $\tau_{m,n}$ are the propagation delays between microphone m and sources l and n respectively, $a_{m,l}$ and $a_{m,n}$ are the mixing coefficients including attenuation factors (due to propagation effects) between microphone m and sources l and n respectively.

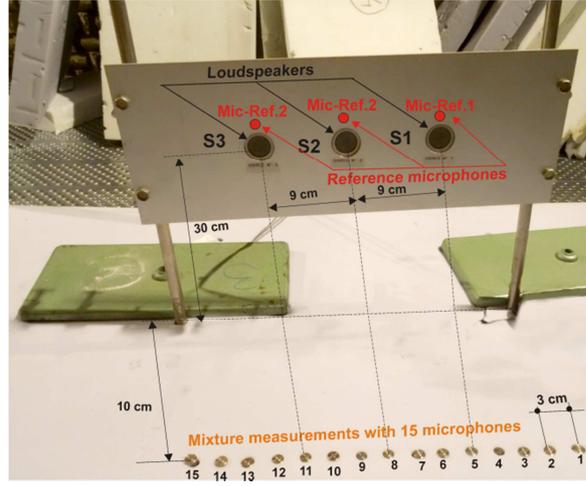


Fig. 2. Experimental set-up in an anechoic chamber to assess the PSDs Bayesian separation approach.

For the three scenarios considered in this article, the objective is to separate the PSDs due to \mathcal{S}_1 , \mathcal{S}_2 and \mathcal{S}_3 starting from, the PSDs of the mixtures $x_m^{1,2}(t)$, $x_m^{1,3}(t)$ and $x_m^{2,3}(t)$ (Eqs. (36), (37)) measured by the microphones of the array. In order to simplify the reading of the article, the following notations are adopted; 1) \mathcal{S}_n ($n = 1$ to 3) is associated to both to source number n and to its PSD \mathcal{S}_n ; 2) $(\mathcal{S}_l + \mathcal{S}_n)$ represents the PSDs of the mixture $x^{l,n}(t)$ ($l = 1$ to 2 and $n = 2$ to 3 with $l \neq n$) measured from the array.

The PSDs of \mathcal{S}_1 , \mathcal{S}_2 and \mathcal{S}_3 are distributed in a frequency bands (FB) which can be given by this relation:

$$FB_i = fH_{si} - fL_{si} \quad (38)$$

It is the difference between fH_{si} and fL_{si} , the upper and lower cutoff frequencies respectively. The frequency bands of the PSDs of the three acoustic sources considered in this article are limited as follows:

$$\begin{cases} FB_1 = 4128 \text{ Hz} - 3872 \text{ Hz} = 256 \text{ Hz} = 2\Delta f \\ FB_2 = 5000 \text{ Hz} - 3000 \text{ Hz} = 2000 \text{ Hz} = 15.625\Delta f \\ FB_3 = 10000 \text{ Hz} - 2000 \text{ Hz} = 8000 \text{ Hz} = 62.5\Delta f \end{cases}$$

Clearly, the problem of PSDs separation is arduous because the frequency bands of the three sources are included the ones in the others as shown in Fig. 3. Indeed, the FB of \mathcal{S}_1 , is included in the FB of \mathcal{S}_2 , which is itself included in the FB of \mathcal{S}_3 .

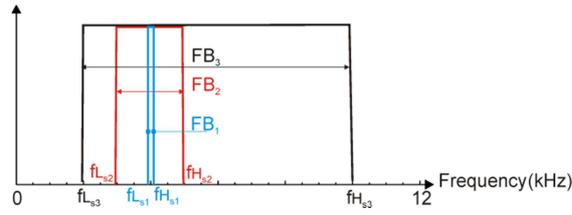


Fig. 3. Sketch showing the width of the frequency bands of the PSDs of the acoustic waves radiated by the three loudspeakers.

The PSDs separation based on a Bayesian approach takes on its full meaning since prior information on the primary source PSDs can be exploited to help their separation from the PSDs of the observations. Indeed, as it is indicated in Fig. 4, we have \mathcal{S}_1 and \mathcal{S}_2 which radiate in frequency ranges much narrower than \mathcal{S}_3 , and \mathcal{S}_1 in a much narrower frequency range than \mathcal{S}_2 . This naturally leads to use a sparsity prior in the frequency domain for \mathcal{S}_1 when it is mixed with \mathcal{S}_2 or \mathcal{S}_3 and for \mathcal{S}_2 when it is mixed with \mathcal{S}_3 to solve the separation problem.

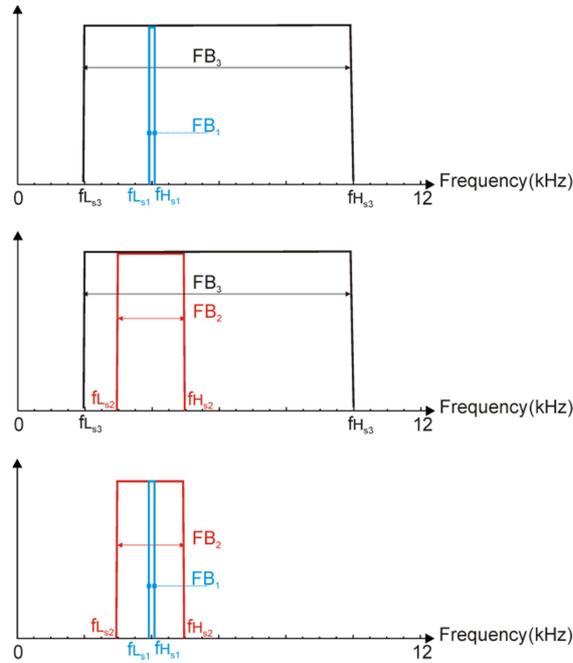


Fig. 4. From top to bottom, comparison of the width of the frequency bands of the PSDs when the sources $(\mathcal{S}_1, \mathcal{S}_3)$, $(\mathcal{S}_2, \mathcal{S}_3)$ and $(\mathcal{S}_1, \mathcal{S}_2)$ are active to justify the use of sparsity prior.

It is pointed out that in the experiment, the microphones of the array are in the far-field of the acoustic sources \mathcal{S}_1 , \mathcal{S}_2 and \mathcal{S}_3 , in the frequency range [1 kHz, 15 kHz] where the three following criteria are satisfied [40]:

$$R \gg \pi D^2 / (2\lambda) \quad (39)$$

$$R \gg D \quad (40)$$

$$R \gg \lambda / (2\pi) \quad (41)$$

where $R = 0.316 \text{ m}$ is the shortest distance from the sources to the measurement positions, λ the wavelength of radiated sound and $D = 0.03 \text{ m}$, is the loudspeaker diameter.

Indeed, the second criterion ((Eq. (40)) is always satisfied since; $R = 0.316 \gg D = 0.03 \text{ m}$. Moreover, the higher frequency $f = 15 \text{ kHz}$ of the emission range, leads to $\lambda_{min} = 0.022 \text{ m}$. Thus, the two criteria (Eqs. (39) - (41)) give:

$$R = 0.316 \gg \frac{\pi D^2}{2\lambda} = (3.14) \frac{0.03^2}{(2)(0.022)} = 0.064 \text{ m} ; R = 0.316 \gg \frac{\lambda}{2\pi} = \frac{0.022}{(2)(3.14)} = 0.0035 \text{ m};$$

In the same way, the lower frequency $f = 1 \text{ kHz}$ of the emission range, leads to $\lambda_{max} = 0.34 \text{ m}$, thus we the two criteria (Eqs. (39) - (41)) that give:

$$R = 0.316 \gg \frac{\pi D^2}{2\lambda} = (3.14) \frac{0.03^2}{(2)(0.34)} = 0.0041 \text{ m} ; R = 0.316 \gg \frac{\lambda}{2\pi} = \frac{0.34}{(2)(3.14)} = 0.054 \text{ m}.$$

This simple verification allows to confirm that the microphones of the array are in the farfield of S_1 , S_2 and S_3 where the sound pressure decreases inversely with the distance.

In contrast, the reference microphones used to build the simulated data are in the nearfield, adjacent to the hydrodynamic field (a region very close to the vibrating surface of the sources). In this area, measurements of the acoustic pressure amplitude give an inaccurate indication of the sound power radiated by the sources [40] due of the presence of interferences between contributing waves from various parts of the sources. Thus, it is difficult to forecast the sound pressure levels in this area because it does not necessarily decrease monotonically at the rate of 6 dB for each doubling of the distance from the sources.

5 TESTS WITH ARTIFICIAL MIXTURES

Before proceeding with the PSDs separation starting from, the PSDs of the measured mixtures, a preliminary step was performed with artificial mixtures of the real source signals. Thus, the signals at the output of the reference microphones *Ref1*, *Ref2* and *Ref3* were filtered in frequency bands [3.8 kHz, 4.2 kHz], [1 kHz, 8 kHz] and [0.3 kHz, 14 kHz] respectively. The results of the filtering constitute the primary source signals $s_1(t)$, $s_2(t)$ and $s_3(t)$. These latter were numerically mixed pairwise according to the model described by (Eq. (36)).

As explained at the conclusion of the previous section, measurements performed in near-field with the reference microphones give an inaccurate indication of the sound power levels

radiated by \mathbf{S}_1 , \mathbf{S}_2 , and \mathbf{S}_3 . Therefore, the magnitudes of the PSDs of the signals at the output of the microphones in farfield cannot be deduced from these measurements. Thus, we have adapted the mixing coefficients in Eq.(36) in order to obtain the same magnitudes that those computed with real mixtures measured with the farfield microphones during the experiment presented in Section 5.

5.1 Analysis of artificial mixtures

The upper plots in Fig.5 show the time series of \mathbf{S}_1 the tonal source, \mathbf{S}_2 the narrowband source and \mathbf{S}_3 the broadband source measured with the microphones *Ref1*, *Ref2* and *Ref3*. The plots in the center present the histogram of \mathbf{S}_1 where two regions are dominant for which the amplitude of \mathbf{S}_1 is either the lowest or the highest. The histograms of \mathbf{S}_2 and \mathbf{S}_3 exhibit a bell shape (center and right graphs) corresponding to a Gaussian distribution. Scatters displayed in the lower plots depict joint distributions of the primary sources. It appears that \mathbf{S}_1 and \mathbf{S}_2 (left graph), \mathbf{S}_1 and \mathbf{S}_3 (center graph) are independent with uniform distribution on a parallelogram. However, the scatter plots for \mathbf{S}_2 and \mathbf{S}_3 show, this time, two independent Gaussian distributions which are rotationally symmetrical (right graph).

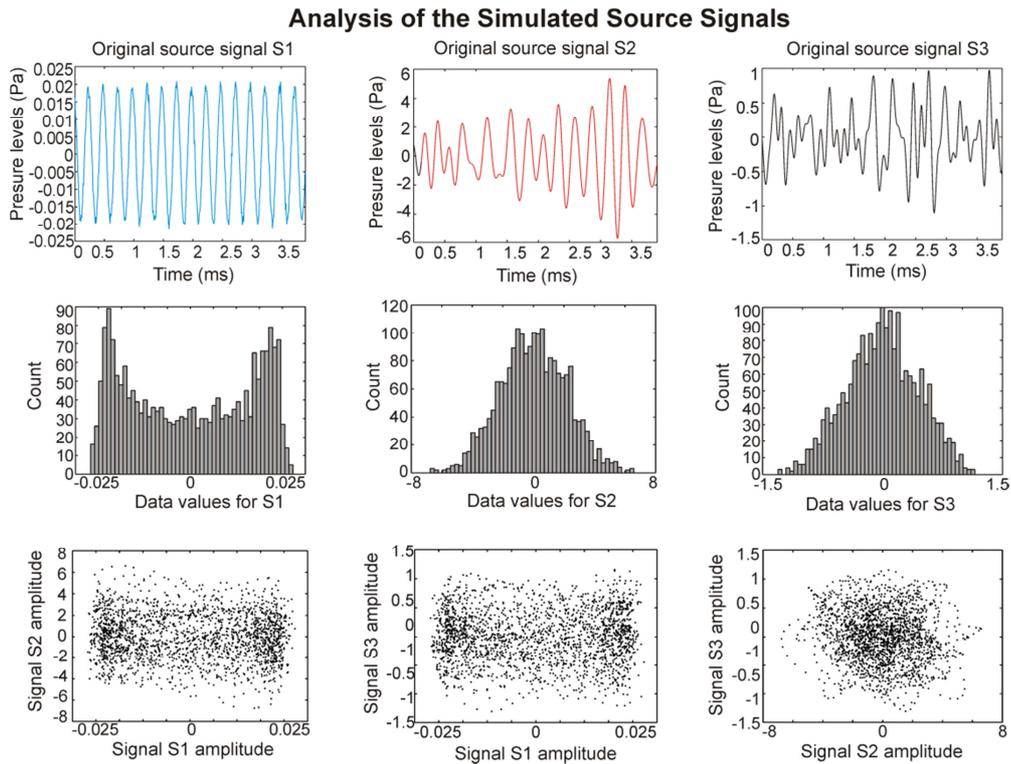


Fig. 5. Primary source signals \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 measured with filtering of *Ref1*, *Ref2*, and *Ref3* shown in Fig.2 (top) - Histograms of \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 (center) - Joint distributions of (\mathbf{S}_1 , \mathbf{S}_2), (\mathbf{S}_1 , \mathbf{S}_3) and (\mathbf{S}_2 , \mathbf{S}_3) (bottom).

The upper graphs in Fig.6 present the time series of the artificial mixtures $x_m^{1,2}(t)$, $x_m^{1,3}(t)$ and $x_m^{2,3}(t)$ (Eq. (36)) obtained when the sources $(\mathbf{S}_1 + \mathbf{S}_2)$, $(\mathbf{S}_1 + \mathbf{S}_3)$ and $(\mathbf{S}_2 + \mathbf{S}_3)$ are respectively active, and measured by the m^{th} ($m = 1$ or 2) microphone of the array. It appears that the histograms of the mixtures tend to be bell-shaped (center graphs in Fig. 6). The scatter plots in the lower graphs show that the mixtures $x_m^{1,2}(t)$ and $x_m^{2,3}(t)$ are no longer independent from each other. However, it is not evident to give the same conclusion for the mixture $x_m^{1,3}(t)$, which looks like an independent Gaussian distribution.

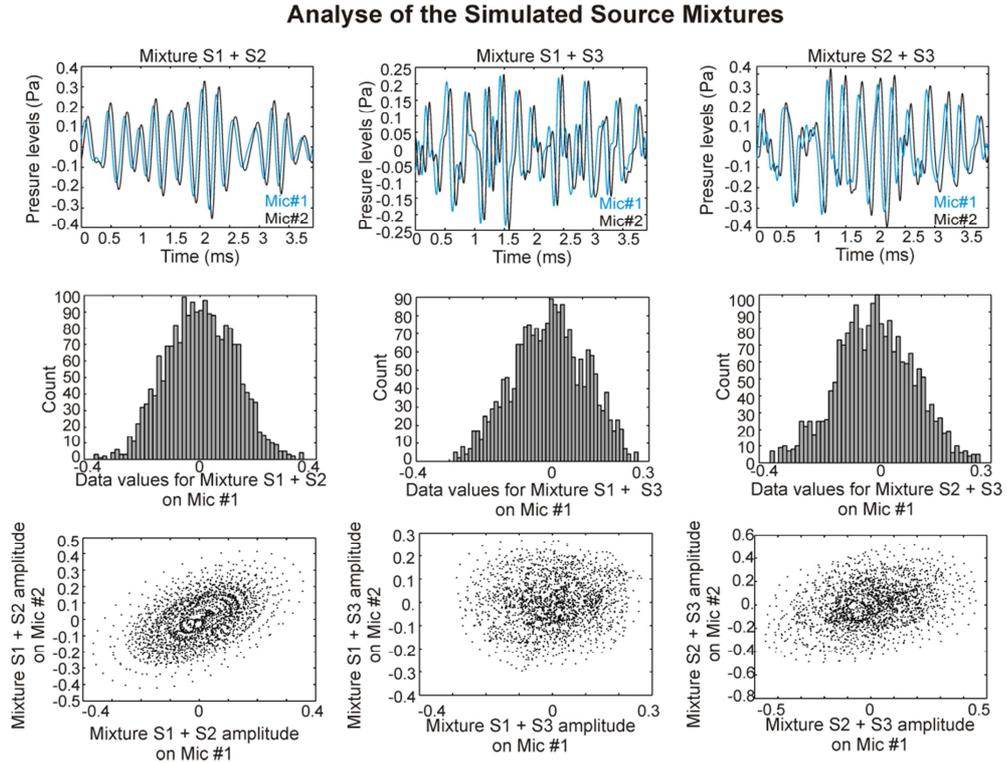


Figure 6. Artificial mixtures $(\mathbf{S}_1 + \mathbf{S}_2)$, $(\mathbf{S}_1 + \mathbf{S}_3)$ and $(\mathbf{S}_2 + \mathbf{S}_3)$ of the primary source signals in the upper plots in Fig.5 (top) - Histogram of the mixtures (center) - Joint distributions of the mixtures (bottom).

The upper graphs in Fig.7 exhibit the PSDs of the primary source signals shown in Fig.5 (upper graphs). In the center, it is superimposed from left to right, the PSDs of $(\mathbf{S}_1, \mathbf{S}_2)$, $(\mathbf{S}_1, \mathbf{S}_3)$ and $(\mathbf{S}_2, \mathbf{S}_3)$, while the lower graphs display the PSDs of the mixtures $x_m^{1,2}(t)$, $x_m^{1,3}(t)$ and $x_m^{2,3}(t)$. These latter will be used as input data of the separation methods proposed in the article to obtain the unmixed PSDs $(\mathbf{S}_1, \mathbf{S}_2)$, $(\mathbf{S}_1, \mathbf{S}_3)$ and $(\mathbf{S}_2, \mathbf{S}_3)$.

PSDs of the Simulated Source Signals

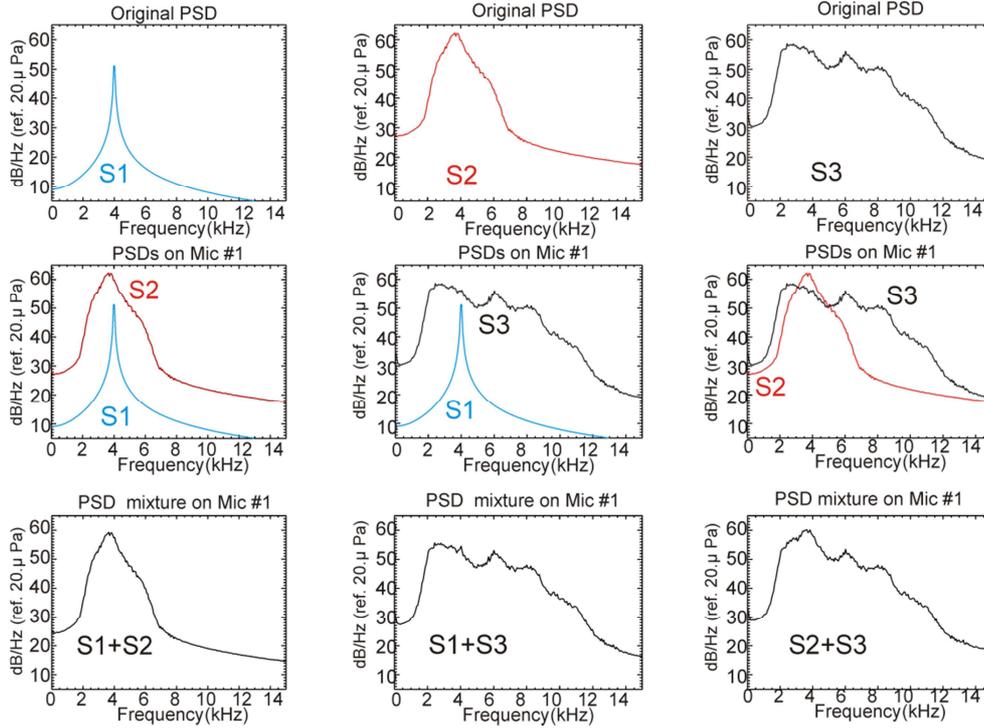


Fig.7. PSDs of primary source signals shown in the upper plots of Fig.5 (top) - Superposition of source signal PSDs (\mathbf{S}_1 , \mathbf{S}_2), (\mathbf{S}_1 , \mathbf{S}_3) and (\mathbf{S}_2 , \mathbf{S}_3) (center) – PSDs of the primary source signals mixed artificially (where $\mathbf{S}_l + \mathbf{S}_n = \text{PSD of } (s_l(t) + s_n(t))$ with $(l, n) = (1,2); (1,3); (2,3)$) (bottom).

6 PRIORS FOR THE SOURCE PSDs AND THE MIXING MATRIX – LIKELIHOOD DESCRIPTION

The approach considered in this study is based on the Joint Maximum A Posterior (JMAP) to solve the separation problem of the PSDs of the sources \mathbf{S} . The JMAP estimator (Eq.(29)) requires to define the priors $p(\mathbf{A})$, $p(\mathbf{S})$ and the likelihood $p(\mathbf{X}|\mathbf{A}, \mathbf{S})$ composing the posterior probability $p(\mathbf{A}, \mathbf{S}|\mathbf{X})$ (Eq. (26)). The forward model for \mathbf{X} (Eq. 18)) is used to describe the observations considered in Section 8.

6.1 PSDs priors

We have different possible priors for the PSDs of the sources \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 , depending on many design choices, and each combination of these choices, giving a different algorithm. We assess the Bayesian separation capability of the PSDs for two kinds of priors described in the next sections.

6.1.1 Gaussian priors for the PSDs of \mathcal{S}

We first consider that there is no specific knowledge of the acoustic PSDs. This leads to adopting a Gaussian distribution as prior law, with zero-mean, and of dispersion law $\sigma_{\mathcal{S}_n}^2$ ($n=1$ to 3) for \mathcal{S}_1 , \mathcal{S}_2 and \mathcal{S}_3 , so that:

$$\mathcal{S}_n \sim \mathcal{N}(\mathcal{S}_n | \mathbf{0}, \sigma_{\mathcal{S}_n}^2 I) \propto \exp \left\{ -\frac{1}{2\sigma_{\mathcal{S}_n}^2} \|\mathcal{S}_n\|_2^2 \right\} \quad (42)$$

where $\|\cdot\|_2$ is the Frobenius norm.

6.1.2 Sparsity prior and Gaussian priors for the PSDs of \mathcal{S}

In this second case one assumes that some knowledge on the acoustic PSDs of \mathcal{S} is available when \mathcal{S}_1 is mixed with \mathcal{S}_2 or \mathcal{S}_3 and when \mathcal{S}_2 is mixed with \mathcal{S}_3 . It is shown in Fig. 4 that the PSDs of \mathcal{S}_1 or \mathcal{S}_2 can be defined, sparse (i.e. has only few nonzero elements in the frequency domain) compared to the PSDs of \mathcal{S}_3 and the PSD of \mathcal{S}_2 sparse compared to the PSD of \mathcal{S}_3 . The sparsity of the PSD can be translated via Laplace density:

$$\mathcal{S}_n \sim \mathcal{L}(\mathcal{S}_n | \alpha_n) \propto \exp \{ -\alpha_n \|\mathcal{S}_n\|_1 \}, \quad n = 1 \text{ or } 2 \quad (43)$$

where $\|\cdot\|_1$ is the L_1 norm.

Moreover, when a sparsity prior is affected to the PSD of \mathcal{S}_1 , a Gaussian one is considered for the PSDs of \mathcal{S}_2 and \mathcal{S}_3 because it is again assumed that we have no specific knowledge on the PSDs of \mathcal{S}_2 and \mathcal{S}_3 . For the same reason, when a sparsity prior is affected to the PSD of \mathcal{S}_2 a Gaussian one is attributed to the PSD of \mathcal{S}_3 .

6.2 Mixing matrix prior

The elements of the mixing matrix \mathbf{A} reflect the coupling between the acoustic sources and the observations achieved by the microphones. Nevertheless, the time signatures of the primary sources \mathcal{S}_1 , \mathcal{S}_2 and \mathcal{S}_3 needed to obtain their PSDs, cannot be accurately measured by using the reference microphones *Ref1*, *Ref2* and *Ref3* for the reasons explained at the end of Section 4. Therefore, the coefficients of the mixing matrix are not available. It is why; we consider here, that the mixing process may be modelled generically because it is not well understood. In order to represent our state of ignorance, a Gaussian distribution of zero-mean with a diagonal covariance matrix $\sigma_A^2 I$ is applied to the mixing matrix. Thus:

$$\mathbf{A} \sim \mathcal{N}(\mathbf{A} | \mathbf{0}, \sigma_A^2 I) \propto \exp \left\{ -\frac{1}{2\sigma_A^2} \|\mathbf{A}\|_2^2 \right\} \quad (44)$$

6.3 Likelihood

The PSD of the noise associated to the experimental measurements can reasonably be represented as a centered Gaussian process:

$$\mathbf{E} \sim \mathcal{N}(\mathbf{E}|\mathbf{0}, \sigma_E^2 I) \propto \exp\left\{-\frac{1}{2\sigma_E^2} \|\mathbf{E}\|_2^2\right\} \quad (45)$$

The choice of the noise covariance indicates that each component of \mathbf{E} in Eq. (18) is contaminated by independent identically distributed random noise. The likelihood can be therefore written in the following form:

$$p(\mathbf{X}|\mathbf{A}, \mathbf{S}, \boldsymbol{\theta}_E) \propto \exp\left\{-\frac{1}{2\sigma_E^2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_2^2\right\} \quad (46)$$

7 JMAP SOLUTIONS FOR THE DATA ANALYSIS

We have defined in Section 3.2 the principle of the optimization method for obtaining generic JMAP solutions $\hat{\mathbf{A}}$ and $\hat{\mathbf{S}}$ for the mixing matrix \mathbf{A} and \mathbf{S} the PSDs of the primary sources. In the following, we examine the possible solutions for the mixing cases defined by Eqs.(36) and (37), by using Eqs.(42), (43), (44) and (45) for the PSDs priors, the matrix prior and the likelihood respectively.

7.1 JMAP - Gaussian priors for the PSDs of \mathbf{S}

In this first situation, we consider that the PSDs of the three acoustic sources \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 have Gaussian priors defined by Eq. (42). By taking into account the prior for \mathbf{A} (Eq. (44)) and the likelihood (Eq.(45), the JMAP estimator becomes (Eq.(33)):

$$J(\mathbf{A}, \mathbf{S}) = \frac{1}{2\sigma_E^2} \|\mathbf{X} - \mathbf{A}_l \mathbf{S}_l - \mathbf{A}_n \mathbf{S}_n\|_2^2 + \frac{1}{2\sigma_A^2} \|\mathbf{A}\|_2^2 + \frac{1}{2\sigma_{S_l}^2} \|\mathbf{S}_l\|_2^2 + \frac{1}{2\sigma_{S_n}^2} \|\mathbf{S}_n\|_2^2 \quad (47)$$

As it is indicated in Section 3.2, the solutions of the cost function $J(\mathbf{A}, \mathbf{S})$ is based on an alternate optimization with respect to \mathbf{A} and then to \mathbf{S} . During the iterative process (Eq. (35)), closed-form solutions are used to compute estimates of \mathbf{A} (when \mathbf{S}_l and \mathbf{S}_n are assumed to be known, l and n defined in Eq. (37)) and of \mathbf{S}_l (when \mathbf{A} and \mathbf{S}_n are assumed to be known) and of \mathbf{S}_n (when \mathbf{A} and \mathbf{S}_l are assumed to be known). This leads to a three step Alternating Least Squares (ALS) algorithm [39]:

$$\begin{cases} \hat{\mathbf{A}} &= \mathbf{X}\hat{\mathbf{S}}^T(\hat{\mathbf{S}}\hat{\mathbf{S}}^T + \lambda_A I)^{-1} \\ \hat{\mathbf{S}}_l &= (\hat{\mathbf{A}}_l^T \hat{\mathbf{A}}_l + \lambda_{S_l} I)^{-1} (\hat{\mathbf{A}}_l^T \mathbf{X} - \hat{\mathbf{A}}_n^T \hat{\mathbf{A}}_l \hat{\mathbf{S}}_n) \\ \hat{\mathbf{S}}_n &= (\hat{\mathbf{A}}_n^T \hat{\mathbf{A}}_n + \lambda_{S_n} I)^{-1} (\hat{\mathbf{A}}_n^T \mathbf{X} - \hat{\mathbf{A}}_l^T \hat{\mathbf{A}}_n \hat{\mathbf{S}}_l) \end{cases} \quad (48)$$

where:

$$\begin{cases} \lambda_A = \frac{\sigma_E^2}{\sigma_A^2} \\ \lambda_{S_l} = \frac{\sigma_E^2}{\sigma_{S_l}^2} \\ \lambda_{S_n} = \frac{\sigma_E^2}{\sigma_{S_n}^2} \end{cases} \quad (49)$$

7.2 JMAP SPARSE - Sparsity prior for the PSD of \mathbf{S}_1 and Gaussian for the PSDs of \mathbf{S}_2 and \mathbf{S}_3

In the second scenario corresponding to the case where \mathbf{S}_1 is mixed with \mathbf{S}_2 or \mathbf{S}_3 (Eq.(36)), one considers that the PSD of \mathbf{S}_1 is sparse and has a Laplace prior Eq. (43). It is also assumed that the PSDs of \mathbf{S}_2 and \mathbf{S}_3 have Gaussian priors (Eq. (34)). By taking into account the prior for \mathbf{A} (Eq.(44)) and the likelihood (Eq.(45)), the JMAP estimator (Eq.(33)) becomes:

$$J(\mathbf{A}, \mathbf{S}) = \frac{1}{2\sigma_E^2} \|\mathbf{X} - \mathbf{A}_l \mathbf{S}_l - \mathbf{A}_1 \mathbf{S}_1\|_2^2 + \frac{1}{2\sigma_A^2} \|\mathbf{A}\|_2^2 + \frac{1}{2\sigma_{S_l}^2} \|\mathbf{S}_l\|_2^2 + \alpha_1 \|\mathbf{S}_1\|_1 \quad (50)$$

where the subscript $l = 2$ or $l = 3$ are used when \mathbf{S}_2 or \mathbf{S}_3 is respectively active.

Thus, according to Eq. (35), the estimates of \mathbf{A} and \mathbf{S} can be obtained using a three step algorithm ALS:

$$\begin{cases} \hat{\mathbf{A}} = \underset{\mathbf{A}}{\operatorname{argmin}} \{ \|\mathbf{X} - \mathbf{A}_l \mathbf{S}_l - \mathbf{A}_1 \mathbf{S}_1\|_2^2 + \lambda_A \|\mathbf{A}\|_2^2 \} \\ \hat{\mathbf{S}}_l = \underset{\mathbf{S}_l}{\operatorname{argmin}} \{ \|\mathbf{X} - \mathbf{A}_l \mathbf{S}_l - \mathbf{A}_1 \mathbf{S}_1\|_2^2 + \lambda_{S_l} \|\mathbf{S}_l\|_2^2 \} \\ \hat{\mathbf{S}}_1 = \underset{\mathbf{S}_1}{\operatorname{argmin}} \{ \|\mathbf{X} - \mathbf{A}_l \mathbf{S}_l - \mathbf{A}_1 \mathbf{S}_1\|_2^2 + \lambda_1 \|\mathbf{S}_1\|_1 \} \end{cases} \quad (51)$$

where:

$$\begin{cases} \lambda_A = \frac{\sigma_E^2}{\sigma_A^2} \\ \lambda_{S_l} = \frac{\sigma_E^2}{\sigma_{S_l}^2} \\ \lambda_1 = \alpha_1 \sigma_E^2 \end{cases} \quad (52)$$

Closed-form solutions are applied to compute estimates of \mathbf{A} (when \mathbf{S}_2 or \mathbf{S}_3 is assumed to be known) and \mathbf{S}_2 or \mathbf{S}_3 (when \mathbf{A} and \mathbf{S}_1 are assumed to be known) during the iterative process:

$$\begin{cases} \widehat{\mathbf{A}} &= X\widehat{\mathbf{S}}^T(\widehat{\mathbf{S}}\widehat{\mathbf{S}}^T + \lambda_A I)^{-1} \\ \widehat{\mathbf{S}}_l &= (\widehat{A}_l^T \widehat{A}_l + \lambda_{S_l} I)^{-1}(\widehat{A}_l^T X - \widehat{A}_n^T \widehat{A}_l \widehat{\mathbf{S}}_n) \end{cases} \quad (53)$$

In contrast, the non-differentiability of the penalty function $\lambda_1 \|\mathbf{S}_1\|_1$ in the objective function $J(\mathbf{S}_1) = \|\mathbf{X} - \mathbf{A}_l \mathbf{S}_l - \mathbf{A}_1 \mathbf{S}_1\|_2^2 + \lambda_1 \|\mathbf{S}_1\|_1$ deriving from Eq.(50) so that there is no closed-form solution for $J(\mathbf{S}_1)$ that can be used in the update rule for obtaining $\widehat{\mathbf{S}}_1$ in the ALS algorithm (Eq. (53)). Therefore, we necessitate to employ an optimization scheme for this step. The L_1 norm regularized least squares LASSO (Least Absolute Shrinkage and Selection Operator) [41] that seeks the minimizer of $J(\mathbf{S}_1)$ is chosen in the optimization process because it provides the best performance in the sense of the sparsity/measurement trade-off.

7.3 JMAP SPARSE - Sparsity prior for the PSD of \mathbf{S}_2 and Gaussian for the PSD of \mathbf{S}_3

This last example is quite similar to the scenario examined in Section 7.2, but here a sparsity prior is applied for the PSD of \mathbf{S}_2 and a Gaussian for the PSD of \mathbf{S}_3 (the subscripts 1 and l are replaced with 2 and 3 respectively in Eqs. (49), (50) and (51)). Again, the non-differentiability of the penalty function $\lambda_2 \|\mathbf{S}_2\|_1$ in the objective function $J(\mathbf{S}_2) = \|\mathbf{X} - \mathbf{A}_2 \mathbf{S}_2 - \mathbf{A}_3 \mathbf{S}_3\|_2^2 + \lambda_2 \|\mathbf{S}_2\|_1$, prevents to have closed-form solutions for $J(\mathbf{S}_2)$ that can be used in the update rule for the solution $\widehat{\mathbf{S}}_2$. This step is again done with regularized least squares LASSO.

7.4 Initialization of ALS algorithm

The ALS solutions given by Eqs. (47) and (50) depends on the parameters λ_A and λ_S in the Eqs.(48) and (51) initialized to positive, arbitrary values for the three scenarios presented in Section 8. However, the determination of those parameters is not easy, in this work we tried to obtain optimal the values of λ_A (in a range $\lambda_{A_{\min}} \leq \lambda_A \leq \lambda_{A_{\max}}$) and λ_S (in a range $\lambda_{S_{\min}} \leq \lambda_S \leq \lambda_{S_{\max}}$) by fixing the optimality criterion to be a good separation of the PSDs of the acoustic sources.

As in all iterative algorithms, the initialization step is needed in the ALS algorithm (Eqs (48) and (51)). A good initialization can improve the speed and accuracy of the algorithm, as it can produce faster convergence an improved minimum. We have achieved satisfactory results by initializing \mathbf{A} with ones values, λ_A and λ_S in the following way:

1. *Scenario # 1 – Gaussian priors for the PSDs of \mathbf{S}*

The PSDs of \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 are initialized with the signal PSDs of the signals at the outputs of the three first microphones of the array and the parameter values of λ_A and λ_S obtained during the ALS iterative processing are:

- $\lambda_A = 100$ is kept with the same value for the two other scenarios;
 - $\lambda_{S_1} = \lambda_{S_2} = \lambda_{S_3} = 0.0001$ when the PSDs have Gaussian priors of equal variances;
 - $\lambda_{S_l} = 0.1$ and $\lambda_{S_n} = 0.4$ (the subscripts $l = 1$; $n = 2$ or 3 , Eq. (37)) when the PSDs have Gaussian priors of non equal variances.
2. *Scenario # 2 – Sparsity prior for the PSD of \mathbf{S}_1 and Gaussian priors for the PSDs of \mathbf{S}_2 and \mathbf{S}_3*

The PSD of \mathbf{S}_1 is initialized with ones everywhere, the PSDs of \mathbf{S}_2 and \mathbf{S}_3 directly by using the PSDs of the signals at the outputs of the two first microphones of the array and the parameter values λ_S obtained during the ALS iterative process are:

- $\lambda_{S_l} = 0.1$ and $\lambda_{S_n} = 0.35$ (the subscripts $l = 1$; $n = 2$ or 3).
3. *Scenario # 3 – Sparsity prior for the PSD of \mathbf{S}_2 and Gaussian one for the PSD of \mathbf{S}_3*

The PSD of \mathbf{S}_2 is initialized with ones everywhere and the PSD of \mathbf{S}_3 directly by using the PSD of the signal at the output of the first microphone of the array and the parameter values λ_S obtained during the ALS iterative process are:

- $\lambda_{S_l} = 0.1$ and $\lambda_{S_n} = 0.35$ (the subscripts $l = 2$; $n = 3$).

8 SEPARATION RESULTS

The PSDs Bayesian separation algorithms presented in Section 7 are applied here on the data recorded during the experiment described in Section 4. In the first part, we examine the separation results of the PSDs of the primary source signals obtained starting from \mathbf{X} (Eq. (18)), the vector of the PSDs of the observations, mixed artificially (see Section 5). In the second part, the results are analyzed when the vector \mathbf{X} is computed with the mixtures measured with the array of microphones (see Section 4). For both kinds of trials:

- the mixing matrix (Eq. (44)) and the likelihood (Eq. (45)) have Gaussian priors;

- all PSDs priors are first considered Gaussian, then sparse and Gaussian, when the observations are mixed artificially, but the priors are only sparse and Gaussian when the observations are measured.

8.1 Separation results obtained from PSDs of observations mixed artificially

8.1.1 Results based on Gaussian priors for the PSDs of \mathcal{S}

8.1.1.1 Gaussian priors of same variance

In this scenario, one considers that the PSDs of the primary sources have all a Gaussian prior of same the variance ($\sigma_{S_1}^2 = \sigma_{S_2}^2 = \sigma_{S_3}^2$ in Eq. (47)). As specified in Sec. 7.4, we have for $\lambda_A = 100$ and $\lambda_{S_1} = \lambda_{S_2} = \lambda_{S_3} = 0.0001$.

The PSDs of the observations mixed artificially pairwise shown in the lower graphs in Fig.7, are unmixed in the right part in Fig.8 based on Eqs.(47), (48) and (49). We observe that $\widehat{\mathcal{S}}_1$ (upper and center plots) and $\widehat{\mathcal{S}}_2$ (bottom plot) are very noisy. By comparing the solutions ($\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2$) and ($\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_3$) to the original source PSDs ($\mathcal{S}_1, \mathcal{S}_2$) and ($\mathcal{S}_1, \mathcal{S}_3$) (left plots in Fig.8), we realize that it is non obvious to conclude that $\widehat{\mathcal{S}}_1$ characterizes the PSD of a tone source at 4kHz. In contrast, $\widehat{\mathcal{S}}_2$ and $\widehat{\mathcal{S}}_3$ are very similar than the original PSDs \mathcal{S}_2 and \mathcal{S}_3 . Finally, separation results for the PSDs of the mixture ($\mathcal{S}_2 + \mathcal{S}_3$) show that $\widehat{\mathcal{S}}_3$ reproduces the spectral shape of the PSD of \mathcal{S}_3 , but the same conclusion cannot be applied for $\widehat{\mathcal{S}}_2$ which does not clearly exhibit that the PSD of \mathcal{S}_2 is spread in the frequency range [2 kHz, 5 kHz].

8.1.1.2 Gaussian priors of not equal variances

This time the variances of the Gaussian priors considered for the PSDs of the primary sources are not equal (*i.e.* $\sigma_{S_l}^2 \neq \sigma_{S_n}^2$ in Eq. (49)). As defined in Sec. 7.4, $\lambda_{S_l} = 0.1$ and $\lambda_{S_n} = 0.4$ (the subscripts l, n are defined in Eq. (37) and λ_{S_l} and λ_{S_n} in Eq. (49)). The separation results of the PSDs based on Eqs.(47), (48) and (49) are presented in Fig. 9 (right plots). The results seem better compared to the those shown in Fig.8 (right plots), obtained when the variances of the Gaussian priors were considered equal. It clearly appears that the background noise has decreased, so that the DSP $\widehat{\mathcal{S}}_1$ of tone source at 4 kHz appears more clearly both when \mathcal{S}_1 is mixed with \mathcal{S}_2 or \mathcal{S}_3 . In contrast, the separation of the PSD of the mixture ($\mathcal{S}_2 + \mathcal{S}_3$) is again, not totally well achieved. Indeed, it is not obvious to conclude that $\widehat{\mathcal{S}}_2$ characterizes of a narrow band in the frequency range [3kHz, 5kHz] source comparing to $\widehat{\mathcal{S}}_3$ which exhibits a broadband source.

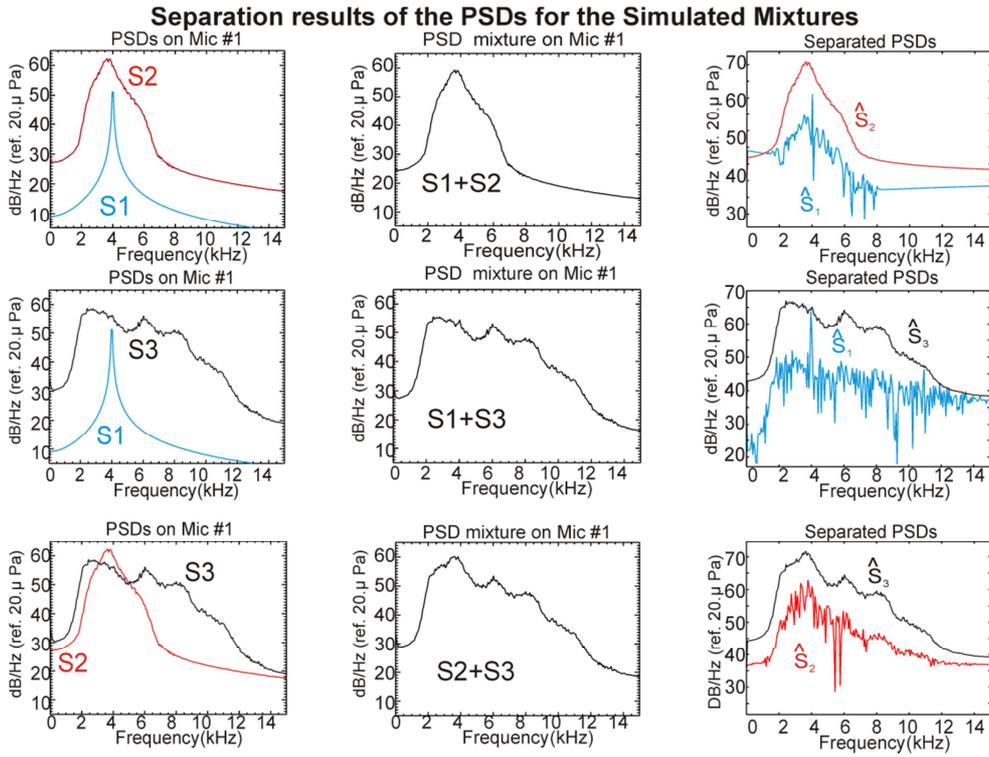


Fig.8. Separation results when the observations are mixed artificially – Left plots: PSDs (S_1 , S_2), (S_1 , S_3) and (S_2 , S_3) of the primary sources S_1 , S_2 and S_3 – Center plots PSDs ($S_l + S_n$) of the mixtures $x_1^{l,n}(t)$ in Eq. (36) ($l = 1$ to 2 ; $n = 2$ to 3 with $l \neq n$) measured with the 1^{st} microphone of the array. Right plots: Unmixed PSDs (S_1 , S_2), (S_1 , S_3) and (S_2 , S_3) by using Gaussian priors of equal variances for the PSDs of S_1 , S_2 and S_3 .

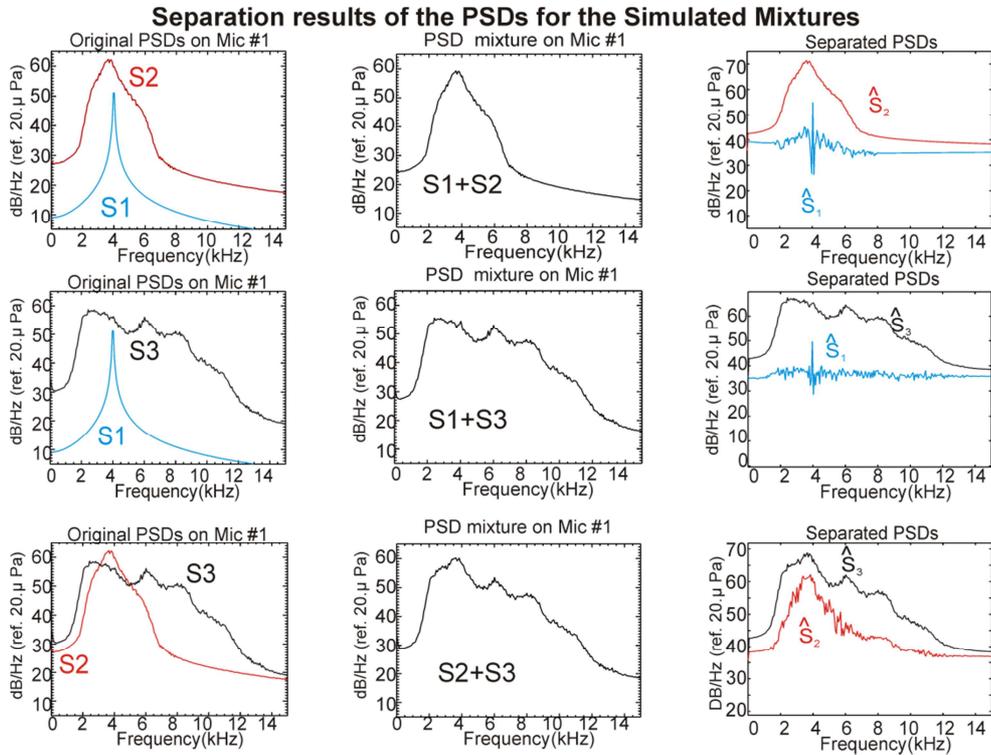


Fig.9. Separation results when the observations are mixed artificially – Left plots: PSDs (\mathbf{S}_1 , \mathbf{S}_2), (\mathbf{S}_1 , \mathbf{S}_3) and (\mathbf{S}_2 , \mathbf{S}_3) of the primary sources \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 – Center plots PSDs ($\mathbf{S}_l + \mathbf{S}_n$) of the mixtures $x_1^{l,n}(t)$ in Eq. (36) ($l = 1$ to 2 and $n = 2$ to 3 with $l \neq n$) measured with the 1st microphone of the array – Right plots: Unmixed PSDs (\mathbf{S}_1 , \mathbf{S}_2), (\mathbf{S}_1 , \mathbf{S}_3) and (\mathbf{S}_2 , \mathbf{S}_3) using Gaussian priors of non-equal variances for the PSDs of \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 .

8.1.2 Results with sparsity and Gaussian priors for the PSDs of S

In this Section, we are again interested to separate the PSDs of the primary source signals which are mixed in the PSDs of the observations shown in the lower plots in Fig.7. A sparsity prior is affected at the PSD of \mathbf{S}_1 and a Gaussian one at the PSD of \mathbf{S}_2 or \mathbf{S}_3 . This time, the separation is based on Eqs. (50), (51) and (52). As defined in Section 7.4, we have: $\lambda_A = 100$, $\lambda_{S_l} = 0.1$ and $\lambda_{S_n} = 0.35$ (the subscripts $l = 1$, $n = 2$ or 3). The separation of the PSDs of the mixtures ($\mathbf{S}_1 + \mathbf{S}_2$) and ($\mathbf{S}_1 + \mathbf{S}_3$) is presented in the right part in Fig. 10. The interest of sparsity prior, clearly appears compared to the results obtained with Gaussian priors plotted in the right part in Figs. 8 and 9. Indeed, $\hat{\mathbf{S}}_1$ shows only a peak at 4 kHz for the tonal source. The characteristics (i.e. shape and frequency support) of the narrowband and broadband sources are again well found with $\hat{\mathbf{S}}_2$ and $\hat{\mathbf{S}}_3$ respectively.

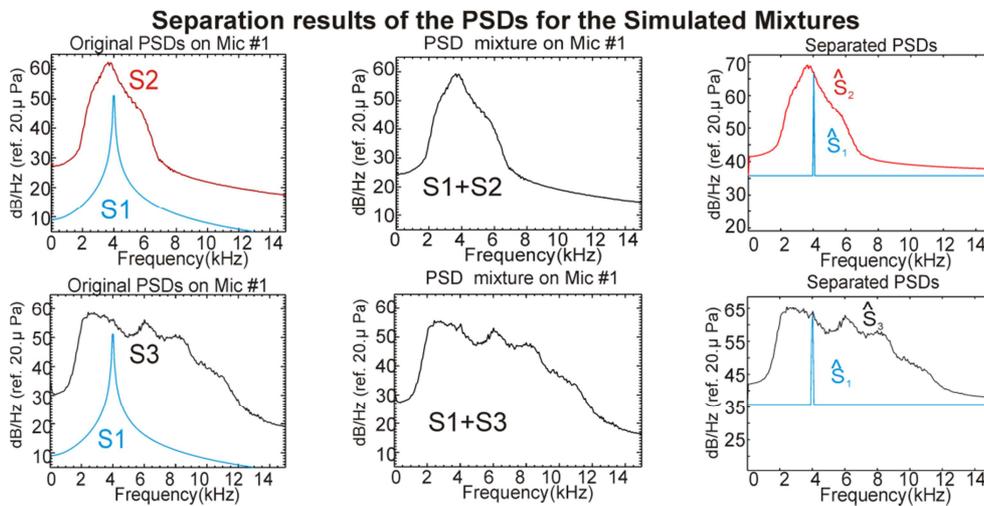


Fig.10. Separation results when the observations are mixed artificially – Left plots: PSDs (\mathbf{S}_1 , \mathbf{S}_2) and (\mathbf{S}_1 , \mathbf{S}_3) of the primary sources \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 – Center plots PSDs ($\mathbf{S}_l + \mathbf{S}_n$) of the mixtures $x_1^{l,n}(t)$ in Eq. (36) ($l = 1$ and $n = 2$ or 3) measured with the 1st microphone of the array – Right plots: Unmixed PSDs (\mathbf{S}_1 , \mathbf{S}_2) and (\mathbf{S}_1 , \mathbf{S}_3) using a Laplace prior for the PSD of \mathbf{S}_1 and a Gaussian priors for the PSDs of \mathbf{S}_2 and \mathbf{S}_3 .

The separation of the PSDs of the mixture ($\mathbf{S}_2 + \mathbf{S}_3$) presented in the right part in Fig. 11 is obtained for $\lambda_A = 100$, $\lambda_{S_l} = 0.1$ and $\lambda_{S_n} = 0.35$ (the subscripts $l = 2, n = 3$). A sparsity prior is affected at the PSD of \mathbf{S}_2 and a Gaussian one at the PSD of \mathbf{S}_3 . Again, the interest of sparsity prior appears clearly. Indeed, \mathbf{S}_2 and \mathbf{S}_3 are well separated with $\hat{\mathbf{S}}_2$ and $\hat{\mathbf{S}}_3$. It was not really the case with the separation results shown in Figs.8 and 9 where Gaussian prior was affected at the PSDs of \mathbf{S}_2 and \mathbf{S}_3 .

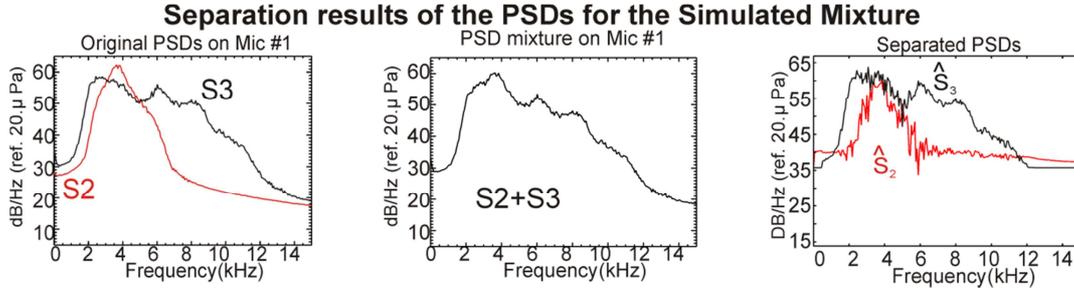


Fig.11. Separation results when the observations are mixed artificially – Left plots: PSDs ($\mathbf{S}_2, \mathbf{S}_3$) of the primary sources \mathbf{S}_2 and \mathbf{S}_3 – Center plots PSDs ($\mathbf{S}_1 + \mathbf{S}_2$) of the mixtures $x_1^{1,2}(t)$ (Eq.(36)) measured with the 1st microphone of the array – Right plots: Unmixed PSDs ($\mathbf{S}_2, \mathbf{S}_3$) using a Laplace prior for the PSD of \mathbf{S}_2 and a Gaussian prior for the PSD of \mathbf{S}_3 .

8.1.3 Evaluation of the results

In order to evaluate Bayesian algorithms, applied to the observations computed with artificial mixtures, we use the signal-to-distortion ratio (SDR) defined in [30], expressed in decibels based on Eq. (8), as follows:

$$SDR[dB] = 10 \log_{10} \frac{\sum_k \sum_{n=1}^N (\mathbf{S}_n(k))^2}{\sum_k \sum_{n=1}^N (\hat{\mathbf{S}}_n(k) - \mathbf{S}_n(k))^2} \quad (54)$$

where \mathbf{S}_n and $\hat{\mathbf{S}}_n$ are the true and separated PSDs referring to the n^{th} source respectively, and the index k stands for the frequency index. \mathbf{S}_n ($n = 1, \dots, N$). The PSDs of the true signals are obtained in a similar manner described at the end of Section 2.3, but in Eq. (19), \tilde{x}_m^ℓ is substituted with \tilde{s}_n^ℓ to give:

$$\mathbf{S}_n = \frac{2}{n_d \Delta t} \sum_{\ell=1}^{n_d} |\tilde{s}_n^\ell|^2, n = 1, \dots, N \quad (55)$$

In each case in Table 1, the best results are achieved when Laplace and Gaussian priors are affected to the PSDs, compared to the cases where only Gaussian priors are used.

It also appears that sparsity plays a role in the quality of the separation measured with SDR. Indeed, SRD is maximum for the separation of the PSDs ($\mathbf{S}_1 + \mathbf{S}_3$), then weakly decreases

for the separation of the PSDs of ($\mathbf{S}_1 + \mathbf{S}_2$) and is minimum for the separation of the PSDs of ($\mathbf{S}_2 + \mathbf{S}_3$).

Finally, one notices that the Gaussian priors with the same variances give higher SRDs than when Gaussian priors with non equal variances are used.

		Mixtures		
		$\mathbf{S}_i + \mathbf{S}_j$ ($i = 1; j = 2$)	$\mathbf{S}_i + \mathbf{S}_j$ ($i = 1; j = 3$)	$\mathbf{S}_i + \mathbf{S}_j$ ($i = 2; j = 3$)
Source priors	Gaussian for \mathbf{S}_i and \mathbf{S}_j with $\sigma_{S_i}^2 = \sigma_{S_j}^2$	SDR = -4.40725	SDR = -2.16497	SDR = -0.926623
	Gaussian for \mathbf{S}_i and \mathbf{S}_j with $\sigma_{S_i}^2 \neq \sigma_{S_j}^2$	SDR = -7.48531	SDR = -8.58410	SDR = -0.273513
	Laplace for \mathbf{S}_i and Gaussian for \mathbf{S}_j	SDR = 10.8997	SDR = 12.5466	SDR = 1.90496

Table 1. Source separation performance measured with SDR

It is important to be able to measure how the PSDs are sparse or dense, one versus the other to explain the results obtained with sparsity priors. This can be done by using a sparseness measure based on the ℓ_0 norm [42], that can reliably quantify the sparseness of the PSD \mathbf{S}_i of the i^{th} source, compared to the length of the frequency band FB_j (Eq.38)) of the PSD \mathbf{S}_j of the j^{th} source. Indeed, consider the PSD of the i^{th} source computed for discrete frequencies f_l ($l = 0, 1, \dots, Q - 1$):

$$\mathbf{S}_i = [S_i(f_0), S_i(f_1), \dots, S_i(f_{Q-1})] \neq 0 \quad (56)$$

Then, define the function:

$$f(S_i(f_l)) = \begin{cases} 1 & \text{for } S_i(f_l) \neq 0 \\ 0 & \text{for } S_i(f_l) = 0 \end{cases} \quad (57)$$

The norm ℓ_0 of \mathbf{S}_i is:

$$\|\mathbf{S}_i\|_0 = \sum_{l=0}^{Q-1} f(S_i(f_l)) \quad (58)$$

Basically, the ℓ_0 norm of the vector \mathbf{S}_i is equal to the number of its non-zero components. For $\mathbf{S}_i \neq 0$, we have always:

$$0 \leq \|\mathbf{S}_i\|_0 \leq Q \quad (59)$$

The most obvious sparseness measure based on the ℓ_0 norm is defined as:

$$\zeta_0(\mathbf{S}_i) = \frac{Q}{Q-1} \left(1 - \frac{\|\mathbf{S}_i\|_0}{Q}\right) \quad (60)$$

We see that closer the measure $\zeta_0(\mathbf{S}_i)$ is to 1, the sparser is the \mathbf{S}_i . On the contrary, the closer the measure is to 0, the denser or less sparse is the \mathbf{S}_i .

One can deduce the ℓ_0 norm of the PSDs \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 by basing on the Eqs. (57) and (58) and by using the frequency bands of the PSDs of the three acoustic sources presented after Eq. (38), thus:

$$\begin{cases} \|\mathbf{S}_1\|_0 = 2 \\ \|\mathbf{S}_2\|_0 = 15.625 \\ \|\mathbf{S}_3\|_0 = 62.5 \end{cases} \quad (61)$$

Table 2 shows the comparison of the sparseness of the PSD \mathbf{S}_i compared to the length of the frequency band FB_j (Eq.38) of the PSD \mathbf{S}_j obtained with Eq. (60) but where Q is substituted with $Q_{ij} = \frac{FB_j}{\Delta f}$ (the index i stands for the PSD of i^{th} source whose we want to compare the sparsity versus the wideness of the frequency band of the PSD of the source of index j). As it was foreseeable, the measure of the sparseness $\zeta_0^{FB_3}(\mathbf{S}_1)$ is greater $\zeta_0^{FB_2}(\mathbf{S}_1)$ which itself greater than $\zeta_0^{FB_3}(\mathbf{S}_2)$, respectively obtained for the PSDs of the mixtures $(\mathbf{S}_1 + \mathbf{S}_3)$, $(\mathbf{S}_1 + \mathbf{S}_2)$ and $(\mathbf{S}_2 + \mathbf{S}_3)$.

Laplace prior for the PSD of \mathbf{S}_i and Gaussian one for the PSD of \mathbf{S}_j		
$\mathbf{S}_i + \mathbf{S}_j$ ($i = 1; j = 2$)	$\mathbf{S}_i + \mathbf{S}_j$ ($i = 1; j = 3$)	$\mathbf{S}_i + \mathbf{S}_j$ ($i = 2; j = 3$)
$Q_{12} = \frac{FB_2}{\Delta f} = 15.625$	$Q_{13} = \frac{FB_3}{\Delta f} = 62.5$	$Q_{23} = \frac{FB_3}{\Delta f} = 62.5$
$\zeta_0^{FB_2}(\mathbf{S}_1) = \frac{Q_{12}}{Q_{12} - 1} \left(1 - \frac{\ \mathbf{S}_1\ _0}{Q_{12}}\right)$	$\zeta_0^{FB_3}(\mathbf{S}_1) = \frac{Q_{13}}{Q_{13} - 1} \left(1 - \frac{\ \mathbf{S}_1\ _0}{Q_{13}}\right)$	$\zeta_0^{FB_3}(\mathbf{S}_2) = \frac{Q_{23}}{Q_{23} - 1} \left(1 - \frac{\ \mathbf{S}_2\ _0}{Q_{23}}\right)$
$\zeta_0^{FB_2}(\mathbf{S}_1) = 0.933333$	$\zeta_0^{FB_3}(\mathbf{S}_1) = 0.983607$	$\zeta_0^{FB_3}(\mathbf{S}_2) = 0.754098$

Table 2. Measure of the sparseness of the PSD of the i^{th} source compared to the frequency band FB_j of the PSD of the j^{th} source.

The results depicted in Table 2 tend to show that the mean SDR presented in Table 1, decreases in a similar manner that the measure of the sparseness $\zeta_0(\mathbf{S}_i)$.

8.2 Experimental results

The Bayesian PSDs separation method previously employed in Section 7 on PSDs of the observations artificially mixed is applied now to the PSDs of the mixtures $(\mathbf{S}_1 + \mathbf{S}_2)$, $(\mathbf{S}_1 + \mathbf{S}_3)$ and $(\mathbf{S}_2 + \mathbf{S}_3)$ measured with the microphones of the array during the experiment described in Section 4 (Fig. 2). The reference signals *Ref1*, *Ref2*, and *Ref3* due to the

loudspeakers \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 respectively (upper plots in Fig.12) are not filtered compared to the cases studied with the artificial mixtures (upper plots in Fig.5). The result, is that the histogram (center plots in Fig.12) of \mathbf{S}_1 characterize a Gaussian distribution and not a sine wave as it was previously presented in the medium of Fig. 5. As anticipated, the histograms for \mathbf{S}_2 and \mathbf{S}_3 characterize Gaussian distributions. The scatter plots of the joint distributions displayed in the lower part in Fig.12, are similar to those of independent Gaussian sources, for $(\mathbf{S}_1, \mathbf{S}_2)$, $(\mathbf{S}_1, \mathbf{S}_3)$ and $(\mathbf{S}_2, \mathbf{S}_3)$.

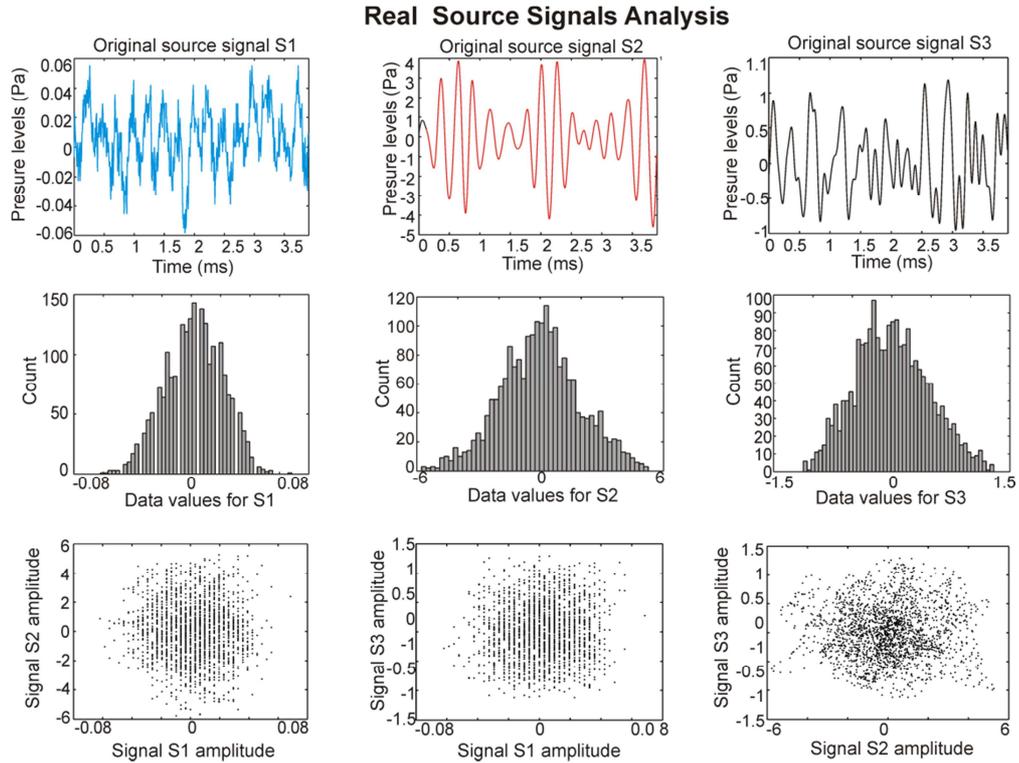


Fig.12. Real data – Source signals of \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 measured without filtering of Ref1, Ref2, and Ref3 shown in Fig.2 (top); Histograms of the primary sources (center); Joint distributions of the primary sources (bottom).

The time series of the mixtures of $(\mathbf{S}_1 + \mathbf{S}_2)$, $(\mathbf{S}_1 + \mathbf{S}_3)$ and $(\mathbf{S}_2 + \mathbf{S}_3)$ measured with the microphones 1 and 2 of the linear array are presented in the upper graphs in Fig.13. The corresponding histograms of the mixtures measured with microphone 1, show Gaussian distributions (center graphs). Like for the tests carried out with artificial mixtures (Fig.6 bottom), it appears that the scatter plots in the lower graphs in Fig. 13 show that the mixtures $(\mathbf{S}_1 + \mathbf{S}_2)$, $(\mathbf{S}_2 + \mathbf{S}_3)$ are no longer independent from each other. In contrast, the mixture $(\mathbf{S}_1 + \mathbf{S}_3)$ looks like independent Gaussian distribution.

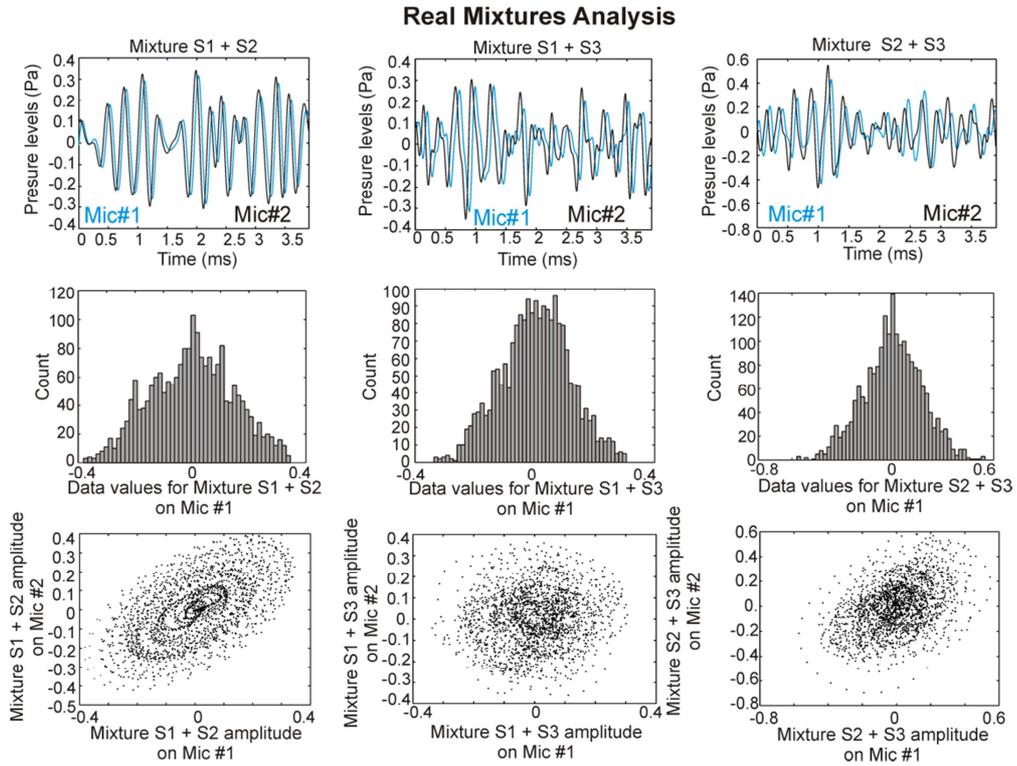


Figure 13. Real mixtures ($\mathcal{S}_1 + \mathcal{S}_2$), ($\mathcal{S}_1 + \mathcal{S}_3$) and ($\mathcal{S}_2 + \mathcal{S}_3$) of the primary source signals shown in the upper plots in Fig.12 (top) – Histogram of mixtures (center), Joint distributions of mixtures (bottom).

The upper plots in Fig.14 exhibit the PSDs of the reference source signals *Ref1*, *Ref2*, and *Ref3* due to \mathcal{S}_1 , radiating a sine wave at 4 kHz, \mathcal{S}_2 , a narrowband wave in the frequency range [3 kHz, 5 kHz] and \mathcal{S}_3 a broadband wave in the interval [2 kHz, 10 kHz] respectively. It is not surprising that the PSD of the signal due to \mathcal{S}_1 is more noisier than its filtered version (upper plots in Fig.7). However, it is not the case for the PSDs of source signals measured with *Ref2* and *Ref3* which are very similar than their filtered versions. The PSDs of (\mathcal{S}_1 and \mathcal{S}_2), (\mathcal{S}_1 and \mathcal{S}_3) and (\mathcal{S}_2 and \mathcal{S}_3) measured with microphone 1 of the array are superimposed in the center of the Fig.14, and the PSDs of mixtures ($\mathcal{S}_1 + \mathcal{S}_2$), ($\mathcal{S}_1 + \mathcal{S}_3$) and ($\mathcal{S}_2 + \mathcal{S}_3$) in the bottom graphs.

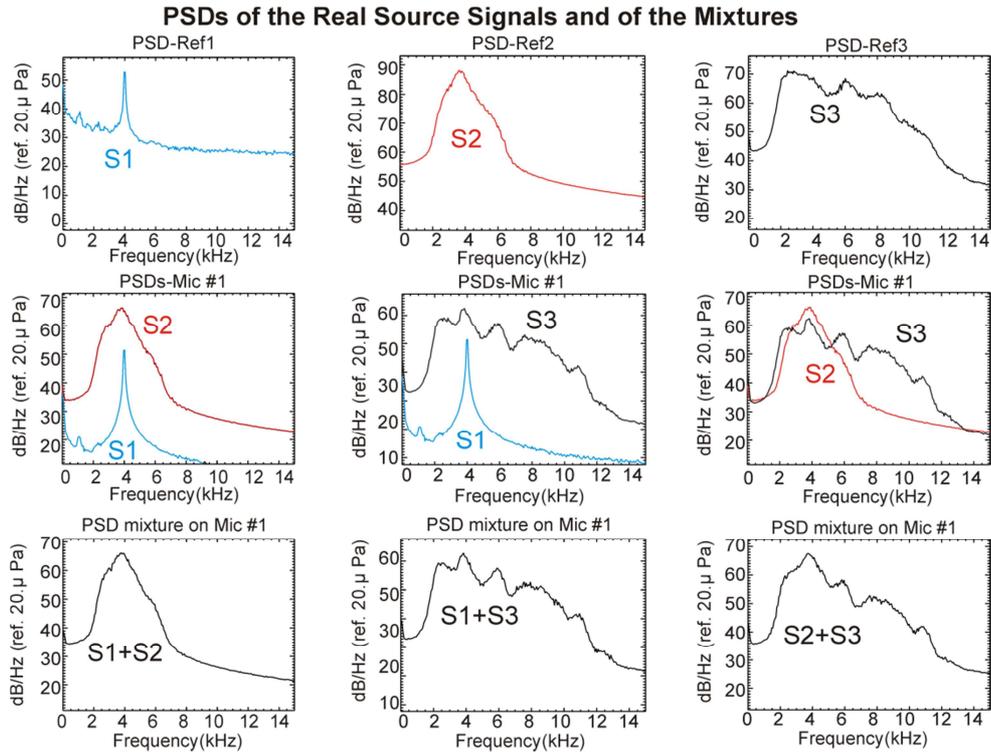


Fig.14. Real mixtures - PSDs of the reference source signals shown in the upper plots in Fig.12 (top) - PSDs (S_1 , S_2), (S_1 , S_3) of the primary sources S_1 , S_2 and S_3 (center) - PSDs of the mixtures ($S_1 + S_2$), ($S_1 + S_3$) and ($S_2 + S_3$) displayed in the upper plots in Fig.13 (bottom).

The results obtained with data mixed artificially have clearly shown the interest to use, both sparsity and Gaussian priors, (Figs. 10 and 11) in the improvement of the separation quality, compared to the scenarios where only Gaussian priors are used (Fig. 8 and 9). This justifies that we limit the application of the Bayesian separation of the PSDs in the cases where sparsity and Gaussian priors are both considered. The Gaussian priors for the mixing matrix (Eq. (44)) and the likelihood (Eq. (45)) are again preferred, for the reasons already explained in Sections 6.2 and 6.3. The separation of the PSDs of the mixtures ($S_1 + S_2$) and ($S_1 + S_3$) and ($S_2 + S_3$) obtained with $\lambda_A = 800$, $\lambda_{S_l} = 0.09$ and $\lambda_{S_n} = 1.9$ ($l = 1$ to 2 and $n = 2$ to 3 with $l \neq n$) is presented in Fig. 15. It clearly appears that the use of sparsity prior is quite relevant in solving the acoustic PSDs demixing problem dealt in the present article.

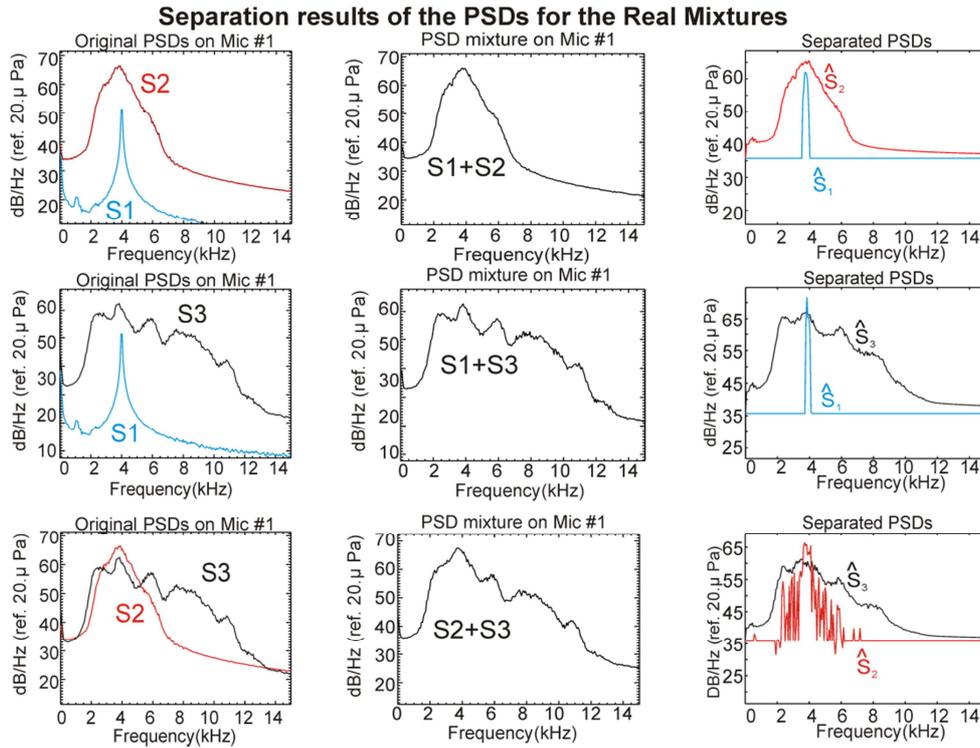


Fig.15. Separation results with real mixtures – Left plots: PSDs (\mathbf{S}_1 ; \mathbf{S}_2), (\mathbf{S}_1 ; \mathbf{S}_3) of the primary sources \mathbf{S}_1 , \mathbf{S}_2 and \mathbf{S}_3 – Center plots PSDs ($\mathbf{S}_l + \mathbf{S}_n$) of the mixtures ($l = 1$ to 2 and $n = 2$ to 3 with $l \neq n$) measured with the 1st microphone of the array – Right plots: Unmixed PSDs i) (\mathbf{S}_1 , \mathbf{S}_2) using Laplace prior for the PSD of \mathbf{S}_1 and a Gaussian one for the PSD of \mathbf{S}_2 (top) ; ii) (\mathbf{S}_1 , \mathbf{S}_3) using Laplace prior for the PSD of \mathbf{S}_1 and a Gaussian one for the PSD of \mathbf{S}_3 (center) and iii) (\mathbf{S}_2 ; \mathbf{S}_3) using Laplace prior for the PSD of \mathbf{S}_2 and a Gaussian one for the PSD of \mathbf{S}_3 (bottom).

9 CONCLUSIONS

In this paper, we considered the Power Spectral Densities (PSDs) separation of the primary source signals from, the PSDs of the measured mixtures with an array of microphones, in a Bayesian framework. The mixtures were obtained during tests performed in an anechoic chamber with three loudspeakers radiating a tonal, narrow-band and broadband acoustic waves. The acoustic sources were mixed pairwise and the separation of their PSDs achieved by considering a Gaussian prior for the PSD of the broadband source, a Laplace or a Gaussian prior for the PSD of the narrowband source and a Laplace prior for the PSD of the tonal source. We compared the separation results based on JMAP method applied to the PSDs of mixtures mixed artificially and on the PSDs obtained with real mixtures for several scenarios. For each test example, first, the expression of the joint posterior law $p(\mathbf{A}, \mathbf{S} | \mathbf{X}, \boldsymbol{\theta})$ was obtained and then optimized via an alternating optimization algorithm.

The main conclusions of this paper are as follows:

- first, the Bayesian approach could be used efficiently on PSDs of mixtures mixed artificially, or obtained with measured mixtures,
- second, the measurement of the quality of the separation of the PSDs provided by the mean Signal-to-Distortion Ratio (SDR) has clearly shown the interest to use, at the same time, sparsity and Gaussian priors for the tests considered in the article. Indeed, this choice has given the best results compared to the tests where only Gaussian priors are used.
- third, it is demonstrated by using a measure of the sparseness of the PSDs based on the ℓ_0 norm, that the quality of the separation given by SDR decreases in a similar way than the degree of sparsity of the PSDs of the acoustic sources.

10 ACKNOWLEDGMENTS

Funding for the present work was provided by Fluid Mechanics Domain and the Aerodynamics, Aeroelasticity and Acoustics Department.

REFERENCES

- [1] H. Ran and D. Mavris, « Preliminary design of a 2d supersonic inlet to maximise total pressure recovery, » in AIAA 5th Aviation, Technology, Integration, and Operations Conference (ATIO), Arlington, VA, 2005.
- [2] J. M. Brookfield and I. A. Waitz, « Trailing-edge blowing for reduction for reduction of turbomachinery fan noise, » Journal of Propulsion and Power, vol. 16, no. 1, pp. 57 – 64, January-February 2000.
- [3] B. D. Mugridge, « The measurement of spinning acoustic modes generated in an axial flow fan, » J. Sound Vib., vol. 10, no. 2, pp. 227 – 246, January-February 1969.
- [4] S. Martens, « Jet noise reduction technology development at GE aircraft engines, » in ICAS 2002 CONGRESS, 2002.
- [5] Y. Li, M. Smith, X. Zhang, Measurement and control of aircraft landing gear broadband noise, Aero. Sci. Technol. 23 (2012) 213-223.

- [6] Y. Tani, Y. Matsuda, A. Doi, Y. Yamashita, S. Aso, and T. Ito, « Experimental study of the morphing flap as a low noise high lift device for aircraft wing, » in 28th International Congress Of The Aeronautical Sciences, 2012.
- [7] C. K. W. Tam, Computational Aeroacoustics A Wave Number Approach. Cambridge University Press, 2012.
- [8] I. L. Griffon, « Aircraft noise modelling and assessment in the IESTA program with focus on engine noise, » in The 22nd International Congress on Sound and Vibration, Florence (Italy), 2015.
- [9] D. Blacodon and G. Elias, « Level estimation of extended acoustic sources using a parametric method, » J. of Aircraft, vol. 41, no. 6, pp. 1360 – 1369, June. 2004.
- [10] D. Blacodon, « Spectral estimation method for noisy data using a noise reference, » Applied Acoustics, vol. 72, no. 1, pp. 11 – 21, January 2011.
- [11] N. Chu, J. Picheral, and A. Mohammad-Djafari, « A robust super-resolution approach with sparsity constraint in acoustic imaging, » Applied Acoustics, vol. 76, pp. 197 – 208, February 2014.
- [12] P. Comon and C. Jutten, Handbook of Blind Source Separation Independent Component Analysis and Applications. Academic Press is an imprint of Elsevier, 2010.
- [13] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization. Nature 401 (6755), 788 – 791 (1999).
- [14] D.D. Lee, H.S. Seung, Algorithms for non-negative matrix factorization, in Proceedings NIPS (2000), pp. 556 – 562.
- [15] P. Smaragdis and J. C. Brown, « Non-negative matrix factorization for polyphonic music transcription, » in Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA), Oct. 2003, pp. 177 – 180.
- [16] N. Bertin, R. Badeau, and G. Richard, « Blind signal decompositions for automatic transcription of polyphonic music : NMF and K-SVD on the benchmark, » in Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP'07), Honolulu, HI, 2007, pp. 65 – 68.

- [17]. A. Ozerov, C. Févotte, Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. *IEEE Trans. Audio Speech Lang. Process.* 18 (3), 550 – 563 (2010).
- [18]. H. Kameoka, T. Yoshioka, M. Hamamura, J. Le Roux, K. Kashino, Statistical model of speech signals based on composite autoregressive system with application to blind source separation, in *Proceedings of International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA)* (2010), pp. 245 – 253
- [19] A. Ozerov, C. Févotte, R. Blouet, J.-L. Durrieu, Multichannel nonnegative tensor factorization with structured constraints for user-guided audio source separation, in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, May 2011, pp. 257 – 260.
- [20] H. Sawada, R. Mukai, S. Araki, S. Makino, Convolutional blind source separation for more than two sources in the frequency domain, in *Proceeding ICASSP* (2004), pp. III-885 – III-888.
- [21] H. Buchner, R. Aichner, W. Kellerman, A generalization of blind source separation algorithms for convolutive mixtures based on second order statistics. *IEEE Trans. Speech and Audio Process.* 13 (1), 120 – 134 (2005).
- [22] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, F. Itakura, Evaluation of blind signal separation method using directivity pattern under reverberant conditions, in *Proceedings ICASSP* (2000), pp. 3140 – 3143.
- [23] A. Hiroe, Solution of permutation problem in frequency domain ICA using multivariate probability density functions, in *Proceedings ICA* (2006), pp. 601 – 608.
- [24] T. Kim, H.T. Attias, S.-Y. Lee, T.-W. Lee, Blind source separation exploiting higher-order frequency dependencies. *IEEE Trans. Audio, Speech, and Lang. Process.* 15 (1), 70 – 79 (2007).
- [25] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, H. Saruwatari, Efficient multichannel nonnegative matrix factorization exploiting rank-1 spatial model, in *Proceedings ICASSP* (2015), pp. 276 – 280.

- [26] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, H. Saruwatari, Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization. *IEEE/ACM Trans. Audio, Speech, and Lang. Process.* 24 (9), 1626 – 1641 (2016).
- [27] C. Févotte and S. J. Godsill, « A Bayesian approach for blind separation of sparse sources, » *IEEE Transactions On Audio, Speech, And Language Processing*, vol. 14, no. 6, pp. 2174 – 2188, Nov. 2006.
- [28] K. H Knuth, « Informed source separation : A Bayesian tutorial », *Proceedings of the 13th European Signal Processing Conference (EUSIPCO 2005)*, Antalya, Turkey.
- [29] H. Snoussi and A. Mohammad-Djafari, « A mean field approximation approach to blind source separation with Lp priors, » in *Signal Processing Conference 2005 13th European*, 2005.
- [30] Vincent, E., Gribonval, R., Févotte, C. : Performance measurement in blind audio source separation. *IEEE Trans. Audio Speech Lang. Process.* 14 (4), 1462 – 1469 (2006).
- Source Separation Methods. Springer Press, 2007.
- [31] C. Chatfield, “The Analysis of Time Series-An Introduction”, Chapman and Hall, London, 1989, pp . 118.
- [32] Julius S. Bendat, Allan G. Piersol, « Random Data: Analysis and Measurement Procedures » Fourth Edition, John Wiley & Sons, Inc,2010, pp. 280.
- [33] P. Comon, « Independent component analysis a new concept ? » *Signal Processing*, vol. 36, pp. 287 – 314, April 1994.
- [34] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, « A blind source separation technique using second-order statistics, » *IEEE Transactions On Signal Processing*, vol. 45, no. 2, pp. 2174 – 2188, Feb. 1997.
- [35] M. S. Pedersen, J. Larsen, U. Kjems, and L. C. Parra, *A Survey of Convolutional Blind Source Separation Methods*. Springer Press, 2007.
- [36] Knuth K.H. 1998. Difficulties applying recent blind source separation techniques to EEG and MEG. In : G.J. Erickson, J.T. Rychert and C.R. Smith (eds.), *Maximum Entropy and Bayesian Methods*, Boise 1997, Kluwer, Dordrecht, pp. 209-222.

- [37] S. Winter, W. Kellermann, H. Sawada, and S. Makino, "MAP-based underdetermined blind source separation of convolutive mixtures by hierarchical clustering and l1-norm minimization," EURASIP Journal on Advances in Signal Processing, Article ID 24717, 2007.
- [38] A. Mohammad-Djafari, "A Bayesian approach to source separation," in Proc. 19th International Workshop on Bayesian Inference and Maximum Entropy Methods (MaxEnt99), Boise, USA, Aug. 19
- [39] A. Mohammad-Djafari, « Bayesian blind deconvolution of images comparing JMAP, EM and BVA with a student-t a priori model, » in Scientific Cooperations International Workshops on Electrical and Computer Engineering Subfields, Koç University, Istanbul/Turkey, 2014.
- [40] D.A. Bies, C.H. Hansen, Engineering Noise Control. Theory and Practice, 4th ed., Spon Press, London, UK, 2009.
- [41] R. Tibshirani, « Regression Shrinkage and Selection via the Lasso, » Journal of the Royal Statistical Society : Series B, vol. 58, no. 1, pp. 267 – 288, 1996.
- [42] Y.Huang, J. Benesty, and J.Chen, Acoustic MIMO Signal Processing. Berlin, Germany:Springer- Verlag, 2006. 7, 10, 37.