Journal of Sound and Vibration 383 (2016) 414–445

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Blind separation of incoherent and spatially disjoint sound sources

Bin Dong [a,*], Jérôme Antoni [a], Antonio Pereira [a,b], Walter Kellermann [c]

[a] Univ Lyon, INSA-Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France
[b] Univ Lyon, École Centrale de Lyon, Laboratoire de Mécanique des Fluides et d'Acoustique, UMR CNRS 5509, F-69134, Écully, France
[c] Institute of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Cauerstaβe 7, D-91058 Erlangen, Germany

ARTICLE INFO

Article history:
Received 20 October 2015
Received in revised form
16 June 2016
Accepted 14 July 2016
Handling Editor: K. Shin
Available online 8 August 2016

Keywords:
Blind source separation
Source identification
Inverse problem
Joint approximate diagonalization

ABSTRACT

Blind separation of sound sources aims at reconstructing the individual sources which contribute to the overall radiation of an acoustical field. The challenge is to reach this goal using distant measurements when all sources are operating concurrently. The working assumption is usually that the sources of interest are incoherent – i.e. statistically orthogonal – so that their separation can be approached by decorrelating a set of simultaneous measurements, which amounts to diagonalizing the cross-spectral matrix. Principal Component Analysis (PCA) is traditionally used to this end. This paper reports two new findings in this context. First, a sufficient condition is established under which "virtual" sources returned by PCA coincide with true sources; it stipulates that the sources of interest should be not only incoherent but also spatially orthogonal. A particular case of this instance is met by spatially disjoint sources – i.e. with non-overlapping support sets. Second, based on this finding, a criterion that enforces both statistical and spatial orthogonality is proposed to blindly separate incoherent sound sources which radiate from disjoint domains. This criterion can be easily incorporated into acoustic imaging algorithms such as beamforming or acoustical holography to identify sound sources of different origins. The proposed methodology is validated on laboratory experiments. In particular, the separation of aeroacoustic sources is demonstrated in a wind tunnel.

1. Introduction

A fundamental issue in noise and vibration engineering is to identify sound sources of different origins. With the development of more and more stringent standards in terms of acoustical quality, especially in the transportation industry, the need of dedicated techniques for localizing, quantifying, and ranking sound sources has become crucial [1]. In this

Abbreviations: 2SO (Joint), Statistical and Spatial Orthogonality; 3S, Supervised Source Separation; BSS, Blind Source Separation; CLT, Central Limit Theorem; CSM, Cross-Spectral Matrix; ESM, Equivalent Source Method; EVD, Eigen-Value Decomposition; HELS, Helmholtz's Equation Least-Squares; ICA, Independent Component Analysis; JAD, Joint Approximate Diagonalization; MCMC, Markov Chain Monte Carlo; NAH, Near-field Acoustical Holography; PCA, Principal Component Analysis; PSD, Power Spectral Density; SNR, Signal-to-Noise Ratio; SO, Statistical Orthogonality (Only); SOI, Source Of Interest; SONAH, Statistically Optimized NAH; STFT, Short Time Fourier Transform; SVD, Singular Value Decomposition

* Corresponding author. Fax: +33 472 43 87 12.
E-mail addresses: bin.dong@insa-lyon.fr, sagittarius1013@gmail.com (B. Dong), jerome.antoni@insa-lyon.fr (J. Antoni), antonio.pereira@ec-lyon.fr (A. Pereira), wk@LNT.de (W. Kellermann).

http://dx.doi.org/10.1016/j.jsv.2016.07.018
0022-460X/

respect, a recurrent challenge is to separate the partial contributions of the different sound sources that contribute to the overall radiated noise, in particular when sources are all active at the same time and have overlapping frequency spectra.

Many methods have been proposed in the past to meet the above requirements and to replace the traditional subsystem masking techniques which are time consuming and prone to influence the operation conditions of the object under study. Among the most popular approaches, acoustic imaging techniques such as beamforming and Near-field Acoustical Holography (NAH) are particularly interesting because they allow contactless measurements recorded by an array of microphones and are rather universal in their principle. Introduction by Maynard, Williams and Lee in the 80's [2,3], NAH has the remarkable capability to indirectly reconstruct sound sources (typically parietal pressure and normal component of particle velocity) with a good spatial resolution and reasonable quantification. Several other sound imaging techniques have been proposed to meet different industrial needs, such as Statistically Optimized NAH (SONAH) [4,5], the Helmholtz's Equation Least-Squares (HELS) [6], the Equivalent Source Method (ESM) [7,8], Bayesian Focalization [9], to name just a few. Reviews of some of these methods can be found for instance in Refs. [10] and [11]. They will be referred to herein as "backpropagation" methods as they all aim at reconstructing the sound source distributions by backpropagating the measured acoustical pressure to the source domain. This implies solving an inverse problem.

Backpropagation methods are good enough to localize and identify the Sources Of Interest (SOIs) when their spacing is large enough as compared to the attainable spatial resolution (e.g. Rayleigh's limit) and when their relative levels are within the available dynamic range. In other situations, when the SOIs are physically very close to each other, slightly overlap in space, and/or exhibit significant differences in level, their visual separation by traditional acoustic imaging techniques may be difficult or even impossible. As a consequence, the reconstructed source distribution still contains a superposition of mixed components that remain to be unraveled.

One way to solve the problem is to exploit the property that sound sources of distinct physical origins can reasonably be assumed mutually independent and then to resort to statistical criteria to achieve their separation. This is the realm of "source separation", whose objectives are of prime interest in practice.

Technically speaking, there are essentially two types of source separation methods found in the literature: Supervised Source Separation (3S) and Blind Source Separation (BSS). 3S methods can separate out any SOI for which an external reference is available. A "reference" is a signal measured simultaneously with the radiated acoustical field and which is perfectly coherent with the SOI (e.g. a vibration signal captured close to the SOI). This implies that it is uncorrelated – i.e. statistically orthogonal – with the other sources in the mixture. Thus, a mean-square-error prediction filter (also called Wiener filter) can be constructed which maps the reference signal to the sound measurements. By definition, the output of the prediction filter is an estimate of the SOI. Other – but theoretically equivalent – implementations are based on the use of partial coherences [12]. Due to its simplicity, the method began to attract attention in the late 70's right after the dual channel analyzers came out. References [13–16] report early applications to acoustic imaging (mainly NAH). The method has been extended later to account for various scenarii such as weakly nonstationary sources [17] and cyclostationary sources [18]. However, 3S methods have fundamental limitations: 1) references must be available, 2) they must be of excellent quality (in a sense to be described shortly), and 3) they must be of sufficient number (at least as many references as SOIs). Requirement (1) is not always fulfilled, in particular due to accessibility constraints or to limited numbers of tracks of the data acquisition system. Requirement (2) is probably the most difficult to attain: it implies the measurement of external signals with theoretically infinite Signal-to-Noise Ratios (SNR), which are fully coherent with the SOI and totally uncorrelated with the other sources. Positions where these conditions are met may not exist at all and, even if they do, their localization would ideally require solving the source identification problem first. Requirement (3) is also a strong one in particular when several sources are to be separated.

In order to alleviate some of these limits, Tomlinson made use of the Principle Component Analysis (PCA) in an attempt to correct a set of non-ideal references [19]. Another elegant solution has been proposed in Ref. [20] which is to replace external references by "numerical" ones returned by a first resolution of the inverse problem. Yet this is likely to succeed only in the case of sources which are initially well separated in space [21–23].

Indeed, early efforts have been spent to avoid the need of any reference at all. This brings us to the second group of source separation methods. Historically, first contributions to the subject are probably due to Price et al. [24] and Otte et al. [25] in the late 80's who proposed to decorrelate a set of measurements in order to force them to comply with the property of references. The resulting uncorrelated signals are then interpreted as "virtual sources". Technically speaking, this amounts to the diagonalization of the Cross-Spectral Matrix (CSM), which may be achieved either by PCA, partial coherences (an implementation of Gram-Schmidt orthogonalization), Cholesky factorization, etc. As pointed out by Price et al., there is no reason that the so separated virtual sources are totally consistent with a physical origin. Reference [26] demonstrates that PCA separation actually holds provided that the SOI is dominating; similarly, virtual sources obtained by partial coherences are meaningful only if the iterative orthogonalization is performed in a pyramidal order where the n-th measurement (used at iteration n) contains no more contributions than n SOIs including the $n-1$ previously extracted ones (at iterations 1,…, $n-1$). Given an arbitrary numbering of the SOIs and of the measurements, this means that the first selected measurement must contain only SOI#1, the second selected measurement – to be orthogonalized with the first one – must contain only SOIs #1 and #2, the third selected measurement – to be orthogonalized with the subspace spanned by the first and the second ones – must contain only SOIs #1, #2 and #3, etc. Needless to say that such a pyramidal order is hardly encountered in practice. A similar limitation was pointed in Ref. [27] in 1976. Although it was early recognized that virtual sources do not match the SOIs in general, the technical literature contains numerous instances of the application of PCA in sound source

identification, sometimes with surprisingly good separation results (see e.g. [28]). Section 3 of this paper establishes a sufficient condition under which virtual sources actually equal the physical sources, which, as far as the authors know, has never been published and which may explain some of the observations found in the literature.

Elaborating on the fact that decorrelation alone is generally not sufficient for separating sources, several new BSS techniques have been proposed during the last decades within the context of Independent Component Analysis (ICA). The paradigm is that statistical independence rather than statistical orthogonality (decorrelation) should be forced, so that a unique solution can be found [29–32]. In other words, ICA replaces the need for references by strong a priori information on the statistical distribution of the SOIs. From the algorithmic point of view, ICA often amounts to diagonalizing the CSM jointly with another matrix comprising non-linear measures of inter-dependence between a set of signals (e.g. higher-order statistics) [33–35], but not only (other sources of "diversity" such as non-stationarity and sparsity can be exploited). Since ICA was first proposed in the field of signal processing in the 80's [36], it has quickly spread in many other disciplines, including acoustics [37–40]. Unfortunately, the direct application of ICA to the separation of sound sources still faces several difficulties, the most critical of which are reminded hereafter.

- First, ICA requires that no more than one SOI has a Gaussian distribution. This is incompatible with most of the noise identification problems encountered in acoustics which are usually solved in the frequency domain: according to the Central Limit Theorem (CLT), the Fourier coefficients quickly tend to be distributed according to a Gaussian distribution even if the original signals are non-Gaussian in the time-domain [41].
- Second, ICA does not embody the spatial dimension related to 1) the sound propagation from the sources to the measurements and to 2) the possible extension of the source shapes (i.e. the fact that sources are spatially distributed) which involve a *spatial* convolution not to be confused with the *time* convolution addressed in "convolutive" BSS methods (e.g. [42]).

The consideration of sound propagation has been investigated in a series of works intending to couple ICA with beamforming [43–50]. However, most of these researches are restricted to separating the directions of arrival of competing sources (localization from the far-field) and do not consider the reconstruction of the spatial distributions of the SOIs. Zhang et al. recently proposed an ICA method based on second-order statistics to blindly separate and reconstruct broadband and spatially distributed sound sources in both the frequency and space domains [51]. It assumes that the mixing coefficients are smooth functions of frequency and makes use of Markov Chain Monte Carlo (MCMC) sampling, a computer intensive method. It is noteworthy that the separation is achieved in the backpropagated domain of the sources, which is surely advantageous to better regularize the inverse problem. In Ref. [52], the authors proposed a method based on the principle of "least spatial entropy" to blindly separate and reconstruct SOIs with compact spatial distribution. Although the performance of the method was found fully satisfactory, its implementation remained quite involved. The question then arises as whether a simpler approach could possibly return comparable separation results under similar working assumptions.

The separation criterion proposed in [52] is quite general and can be applied to separate sound sources no matter whether they overlap in space or not. One way to restrict the problem is to consider the specific case where sound sources are emitted from disjoint regions, an assumption which in spite of being less general still reflects many real-life configurations, as shown in Fig. 1.

The object of this paper is to propose a simple and fast method – "fast" as compared to the method published in Ref. [52] – to blindly separate and reconstruct sound sources, possibly of broadband nature, that radiate from disjoint domains (the SOIs may otherwise overlap both in the time and frequency domains). Note that the problem is not trivial, since the fact that sound sources are spatially disjoint is not necessarily "seen" by the array of microphones placed in the radiated acoustical field. Indeed, the partial acoustical fields radiated by the individual sources on the array will fully overlap in general, as shown in Fig. 1.

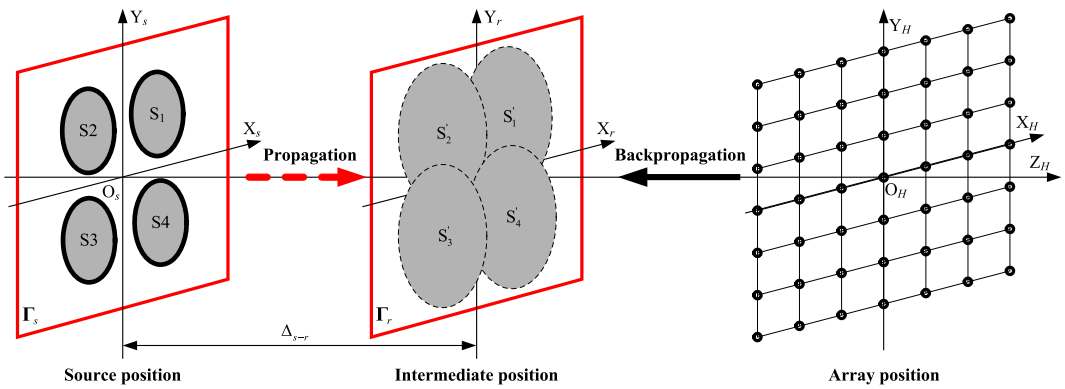

Fig. 1. Schematic illustration of disjoint sources overlapping in the domain of the measurements. The intermediate plane illustrates that the sound fields radiated by disjoint sources will overlap in general.

The rest of the paper is organized as follows. Section 2 addresses the backpropagation of the measurements to the source domain, the resulting mixing model inherent to BSS, and discusses the non-uniqueness of the separation returned by PCA. The condition under which virtual sources returned by PCA actually match the SOIs is then elucidated in Section 3. Next, Section 4 presents a method to separate incoherent and spatially disjoint sound sources by jointly enforcing statistical and spatial orthogonality. The proposed methodology is eventually validated in Section 5 via numerical and laboratory experiments. An illustration is also provided for the separation of aeroacoustic sources in an open-jet anechoic wind tunnel. Conclusions are drawn in Section 6.

2. Blind source separation from decorrelation

As discussed in the previous section, specificities of the acoustical context require a careful formulation of the BSS problem. Since the aim is to separate and reconstruct the SOIs from distant pressure measurements, BSS has to be coupled with a backpropagation procedure (a mathematical operation detailed in Subsection 2.1). In principle, BSS could be performed either before backpropagation, directly on the measured microphone pressures, or after backpropagation, in the space of the reconstructed sources. The second approach is adopted in this paper since it is more amenable to the consideration of spatial properties of the SOIs, as explained in Section 4. Therefore, the BSS problem will be formulated on some expansion coefficients of the sources.

2.1. Backpropagation of measurements to the source domain

Assume that an array of microphones spatially samples the sound field radiated from a set of spatially distributed and mutually incoherent sound sources, referred to as the SOIs. Incoherence means that the SOIs are statistically orthogonal as a consequence of originating from different physical mechanism. The aim is to recover each SOI, which implies their separation and individual reconstruction.

The sound field is considered statistically stationary and ergodic. There is no specific assumption on the probability density functions and the power spectral densities of the SOIs, which may be Gaussian or not, and narrow- or broad-band. A linear mapping is supposed to hold between the source field and the measured pressures, typically characterized by a set of Green functions or transfer functions hereafter denoted as the "propagation operator".

The pressure measurement at the m-th microphone of the array is denoted as $p(\mathbf{r}_m, t)$, where $\mathbf{r}_m$ stands for the position vector of the m-th microphone and t for the time index. The signal $p(\mathbf{r}_m, t)$ is then transformed into a series of "snapshots", $p(\mathbf{r}_m, \omega; \varpi)$, at angular frequency ω, and indexed by ϖ (time position of the snapshots), by application of the Short Time Fourier Transform (STFT). Then, each snapshot $p(\mathbf{r}_m, \omega; \varpi)$ is modeled as

$$p(\mathbf{r}_m, \omega; \varpi) = \int_\Gamma G(\mathbf{r}_m, \mathbf{r}; \omega) s(\mathbf{r}, \omega; \varpi) d\Gamma(\mathbf{r}) + \varepsilon(\omega; \varpi), \tag{1}$$

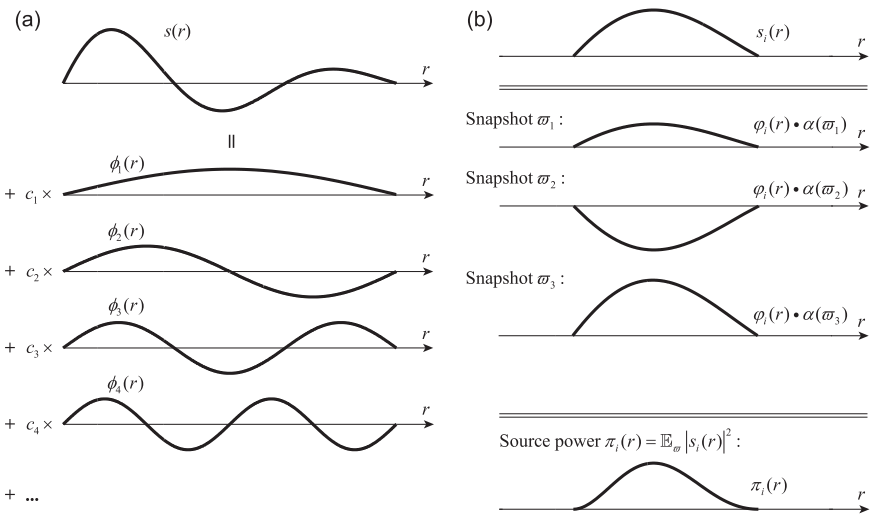

Fig. 2. (a) Expansion of the sound field $s(\mathbf{r})$ (superposition of all sources of interest) onto spatial basis functions $\phi_b(\mathbf{r})$. (b) Expansion of a given source of interest $s_i(\mathbf{r})$ onto spatial mode shapes $\varphi_i(\mathbf{r})$.

with

$$s(\mathbf{r},\omega;\varpi)=\sum_{i=1}^{N_s} s_i(\mathbf{r},\omega;\varpi), \tag{2}$$

where $G(\mathbf{r}_m,\mathbf{r};\omega)$ stands for the propagation operator between the array of microphones and the sound sources $s_i(\mathbf{r},\omega;\varpi)$, $\mathbf{r}$ for the position vector, N_s for the number of active sound sources, Γ for the source domain, and $\boldsymbol{\varepsilon}(\omega;\varpi)$ for "measurement" noise (a combination of instrumentation noise, modeling errors, etc.). Note that Eq. (2) explicitly expresses the total source field as a superposition of N_s sound sources.

In an attempt to reconstruct sound sources from measurements, the total source distribution $s(\mathbf{r},\omega;\varpi)$ is modeled as a weighted summation of spatial basis functions $\phi_b(\mathbf{r},\omega)$,

$$s(\mathbf{r},\omega;\varpi)=\sum_{b=1}^{B}\phi_b(\mathbf{r},\omega)c_b(\omega;\varpi)=\boldsymbol{\phi}(\mathbf{r},\omega)\mathbf{c}(\omega;\varpi), \tag{3}$$

where c_b denotes the expansion coefficient of the b-th spatial basis function $\phi_b(\mathbf{r},\omega)$, $\boldsymbol{\phi}=[\phi_1\ \phi_2\ \cdots\ \phi_B]$, $\mathbf{c}=[c_1\ c_2\ \cdots\ c_B]^T$, and superscript ^{T} indicates the transpose operator. Formulation (3) is left intentionally general so as to accommodate for various linear sound imaging methods which have been presented in the literature. For instance classical NAH uses the discrete Fourier basis in $\boldsymbol{\phi}$ and SONAH an overcomplete set of plane waves, in which cases the expansion coefficients in $\mathbf{c}$ represent values of the spatial spectrum of the sound field; besides, beamforming uses spatial Dirac impulses (e.g. monopoles) for the column of $\boldsymbol{\phi}$ and the expansion coefficients in $\mathbf{c}$ then represent the source strengths. Among other possible choices, the optimal spatial functions proposed in Ref. [9] may be preferred as they achieve a minimum reconstruction error with the number B of basis functions then being equal to the number M of microphones in the array. By means of an example, Fig. 2 (a) illustrates the expansion of a one-dimensional source onto a set of sinusoidal spatial basis functions.

For convenience, let us now consider the discretized version of the problem which, independently of the used discretization scheme, leads to the general formulation

$$\mathbf{p}(\omega;\varpi)=\mathbf{G}(\omega)\boldsymbol{\Phi}(\omega)\mathbf{c}(\omega;\varpi)+\boldsymbol{\varepsilon}(\omega;\varpi), \tag{4}$$

where $\mathbf{p}=[p(\mathbf{r}_1)\ \ p(\mathbf{r}_2)\ \ \cdots\ p(\mathbf{r}_M)]^T$, $\boldsymbol{\Phi}=[\phi(\mathbf{r}_1)^T\ \phi(\mathbf{r}_2)^T\cdots\phi(\mathbf{r}_N)^T]^T$ is an $N\times B$ matrix, and $\mathbf{G}(\omega)$ an $M\times N$ matrix, where N stands for the number of discretization points.

The traditional way to reconstruct the source field $s(\mathbf{r},\omega;\varpi)$ from pressure measurements $p(\mathbf{r}_m,\omega;\varpi))$ is to estimate the unknown vector of coefficients $\mathbf{c}(\omega;\varpi)$ by inverting the system of Eq. (4), that is

$$\hat{\mathbf{c}}(\omega;\varpi)=[\mathbf{G}(\omega)\boldsymbol{\Phi}(\omega)]^{+}\mathbf{p}(\omega;\varpi), \tag{5}$$

where $[\mathbf{G}\boldsymbol{\Phi}]^{+}$ stands for some inverse operator. Eq. (5) reflects the backpropagation operation performed from the measurement to the source domain.

As before, the notation is left sufficiently general to accommodate for different backpropagation methods. In NAH, $[\mathbf{G}\boldsymbol{\Phi}]^{+}$ stands for the pseudo-inverse of the matrix $\mathbf{G}\boldsymbol{\Phi}$ which usually requires careful regularization (e.g. see Ref. [9]). In beamforming, $[\mathbf{G}\boldsymbol{\Phi}]^{+}$ comprises the steering vectors in all scanned directions (i.e. it is made of the conjugate inverse elements of $[\mathbf{G}\boldsymbol{\Phi}]$).

Eq. (5) solves the source reconstruction problem. A further step is now necessary to separate the SOIs, that is to solve Eq. (2) for each of its constituent. As stated in Ref. [52], the SOIs can be decomposed onto spatial mode shapes

$$s_i(\mathbf{r},\omega;\varpi)=\varphi_i(\mathbf{r},\omega)\alpha_i(\omega;\varpi), \tag{6}$$

where φ_i stands for the spatial mode shape corresponding to the i-th source s_i and α_i for its assigned modal coordinate. This is illustrated in Fig. 2(b). Contrary to model (3), Eq. (6) is more amenable to source separation as will be seen shortly. Without loss of generality, coefficient α_i is assumed to be a random variable with zero mean and unit variance (the scale is absorbed in the mode shapes). Since the sound sources are assumed mutually incoherent, they are statistically orthogonal in the sense that

$$\mathbb{E}_{\varpi}\{s_i(\mathbf{r},\omega;\varpi)s_j^{*}(\mathbf{r}',\omega;\varpi)\}=0,\quad i\neq j,\quad \forall(\mathbf{r},\mathbf{r}'), \tag{7}$$

where $\mathbb{E}_{\varpi}$ stands for the averaging operator over snapshots and superscript * for the complex conjugate. Combining Eqs. (6 and 7), one has

$$\mathbb{E}_{\varpi}\{\alpha_i(\omega;\varpi)\alpha_j^{*}(\omega;\varpi)\}=\delta_{ij}, \tag{8}$$

where δ_{ij} stands for the Kronecker delta. Therefore, applying Eqs. (6–8), the power of the i-th source is simply defined as (see Fig. 2(b))

$$\pi_i(\mathbf{r},\omega)=\left|\varphi_i(\mathbf{r},\omega)\right|^2, \tag{9}$$

that is the squared magnitude of the (scaled) mode shape.

Let us now expand the mode shape $\varphi_i(\mathbf{r},\omega)$ onto the spatial basis functions $\phi_b(\mathbf{r},\omega)$,

$$\varphi_i(\mathbf{r},\omega) = \sum_{b=1}^{B} \phi_b(\mathbf{r},\omega) a_{ib}(\omega) = \boldsymbol{\phi}(\mathbf{r},\omega)\mathbf{a}_i(\omega), \tag{10}$$

where a_{ib} stands for the expansion coefficient of the b-th spatial basis function $\phi_b(\mathbf{r},\omega)$ and $\mathbf{a}_i = [a_{i1}\ a_{i2}\ \cdots\ a_{iB}\ ]^T$. Thus combining Eqs. (2), (3), (6), and (10) and dealing with all the N_s sources together, one arrives at

$$\mathbf{c}(\omega;\varpi) = \mathbf{A}(\omega)\boldsymbol{\alpha}(\omega;\varpi), \tag{11}$$

with $\mathbf{A} = [\mathbf{a}_1\ \ \mathbf{a}_2\ \ \cdots\ \mathbf{a}_{N_s}]$ and $\boldsymbol{\alpha} = [\alpha_1\ \alpha_2\ \ \cdots\ \alpha_{N_s}\ ]^T$. Finally, substituting $\hat{\mathbf{c}}$ of Eq. (5) in Eq. (11), the estimated vector $\hat{\mathbf{c}}$ can be rewritten as

$$\hat{\mathbf{c}}(\omega;\varpi) = \mathbf{A}(\omega)\boldsymbol{\alpha}(\omega;\varpi) + \mathbf{n}(\omega;\varpi), \tag{12}$$

where $\mathbf{n}$ embodies the difference between the two representations, $\mathbf{A}\boldsymbol{\alpha}$ and $[\mathbf{G}\boldsymbol{\Phi}]^+\mathbf{p}$, or loosely speaking, the "estimation noise". So far, only the left-hand side of Eq. (12) is available, but the variables on the right side are still unknown, which is the standard description of the BSS problem [32,34]. It is seen that if one succeeds to recover the mixing matrix $\mathbf{A}$ in model (12), the source powers π_i can be individually estimated according to Eqs. (9) and (10). This is the aim of the rest of the paper.

2.2. PCA and virtual sources

Let us now proceed by investigating how to recover the mixing matrix $\mathbf{A}$. Variable ω will be ignored hereafter for the purpose of notational simplification whenever there is no ambiguity. The matrix $\mathbf{A}$ has the Singular Value Decomposition (SVD)

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^H, \tag{13}$$

where, $\mathbf{U}$ and $\mathbf{V}$ are two unitary matrices such that $\mathbf{U}\mathbf{U}^H = \mathbf{I}_B$ and $\mathbf{V}\mathbf{V}^H = \mathbf{I}_{N_s}$, respectively, and $\mathbf{D}$ is a $B \times N_s$ nonnegative diagonal matrix (superscript ^{H} accounts for the Hermitian transpose and $\mathbf{I}$ for the identity matrix). Combining Eqs. (8) and (12 and 13), the CSM of the estimated coefficient vector $\hat{\mathbf{c}}$ can be written as

$$\mathbb{E}_{\varpi}\left\{\hat{\mathbf{c}}\hat{\mathbf{c}}^H\right\} = \mathbf{A}\mathbf{A}^H + \sigma_n^2\mathbf{I} = \mathbf{U}\mathbf{D}^2\mathbf{U}^H + \sigma_n^2\mathbf{I}, \tag{14}$$

where σ_n^2 denotes the variance of the noise $\mathbf{n}$, assuming spatially white random variable. Equation (14) is recognized as the Eigen-Value Decomposition (EVD) of matrix $\mathbb{E}_{\varpi}\{\hat{\mathbf{c}}\hat{\mathbf{c}}^H\}$, which thus returns a unique solution of the two matrices $\mathbf{U}$ and $\mathbf{D}$[53].

This is the basis of the PCA method, which consists in finding a set of uncorrelated sound sources, so called "virtual sources" by diagonalizing the CSM (either of the source expansion coefficients as in Eq. (14) or directly of the pressure measurements) [24,25]. Because it is an orthogonal transformation it conserves energy and the sum of the powers of the virtual sources actually equal the sum of the powers of the SOIs, which is a remarkable property. However, this is not enough to conclude that the virtual sources equal the SOIs.

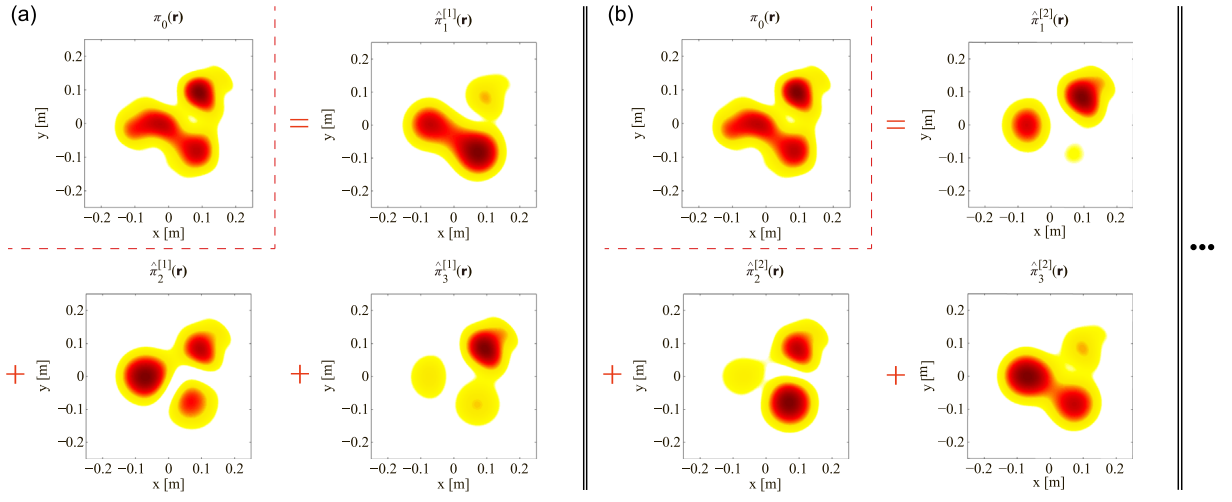

Fig. 3. Schematic illustration of source separation by statistical orthogonalization (e.g. PCA) using two different unitary matrices (a) $\mathbf{V}_1$ and (b) $\mathbf{V}_2$: although the sum of partial powers is conserved, the virtual sources are not uniquely linked to the true sources.

To see this, let us explicitly formulate the partial source power π_i (power contribution from source i). Using Eqs. (9), (10), and (13), one arrives at

$$\pi_i(\mathbf{r}) = \left|\varphi_i(\mathbf{r})\right|^2 = \boldsymbol{\phi}(\mathbf{r})\mathbf{a}_i\mathbf{a}_i^H\boldsymbol{\phi}^H(\mathbf{r}) = \boldsymbol{\phi}(\mathbf{r})\overbrace{\mathbf{UDV}^H\mathbf{e}_i}^{\mathbf{a}_i}\overbrace{\mathbf{e}_i^T\mathbf{VDU}^H}^{\mathbf{a}_i^H}\boldsymbol{\phi}^H(\mathbf{r}), \tag{15}$$

where vector $\mathbf{e}_i$ is the i-th column of the identity matrix. Now, considering virtual sources, $\mathbf{a}_i\mathbf{a}_i^H$ in formula (15) is implicitly replaced by the eigen contribution $d_i\mathbf{u}_i\mathbf{u}_i^H$ (where $\mathbf{u}_i$ is the i-th column of matrix $\mathbf{U}$ and d_i the i-th diagonal element of matrix $\mathbf{D}$), thus leading to

$$\pi_i^{virtual}(\mathbf{r}) = \boldsymbol{\phi}(\mathbf{r})\left(d_i^2\mathbf{u}_i\mathbf{u}_i^H\right)\boldsymbol{\phi}^H(\mathbf{r}) = \boldsymbol{\phi}(\mathbf{r})\mathbf{UDe}_i\mathbf{e}_i^T\mathbf{DU}^H\boldsymbol{\phi}^H(\mathbf{r}). \tag{16}$$

Although by construction the sum of the partial powers is conserved,

$$\sum_{i=1}^{N_s}\pi_i^{virtual}(\mathbf{r}) = \boldsymbol{\phi}(\mathbf{r})\underbrace{\mathbf{UD}^2\mathbf{U}^H}_{\sum_{i=1}^{N_s}\mathbf{a}_i\mathbf{a}_i^H}\boldsymbol{\varphi}^H(\mathbf{r}) = \sum_{i=1}^{N_s}\pi_i(\mathbf{r}), \tag{17}$$

there is no reason why the virtual sources should match the true sources in general. In fact, comparing Eqs. (16) and (15), it is readily seen that

$$\pi_i^{virtual}(\mathbf{r}) \neq \pi_i(\mathbf{r}). \tag{18}$$

The conservation of power reflected by Eq. (17) is actually satisfied by any unitary matrix $\mathbf{V}$ in Eq. (15), thus leading to sound sources with different spatial distributions. This leaves infinity of solutions to the BSS problem, as schematically illustrated in Fig. 3. More specifically, by equating Eqs. (15) and (16), the i-th true source reads

$$\mathbf{s}_i(\mathbf{r}) = \boldsymbol{\phi}(\mathbf{r})\mathbf{UDV}^H\mathbf{e}_i = \underbrace{\boldsymbol{\phi}(\mathbf{r})\mathbf{UD}}_{\left[\mathbf{s}_1^{virtual}\ \cdots\ \mathbf{s}_{N_s}^{virtual}\right]}\underbrace{\mathbf{V}^H\mathbf{e}_i}_{\begin{bmatrix}v_{1i}^*\\ \vdots\\ v_{N_s,i}^*\end{bmatrix}} = \sum_{k=1}^{N_s}v_{ki}^*\mathbf{s}_k^{virtual}(\mathbf{r}), \tag{19}$$

with v_{ki}^* the element (k, i) of matrix $\mathbf{V}^H$, which proves that the true source $\mathbf{s}_i(\mathbf{r})$ is a linear combination of the virtual sources $\mathbf{s}_k^{virtual}(\mathbf{r})$ – and vice-versa.

In conclusion, statistical orthogonality as achieved by PCA is a necessary but not a sufficient condition for source separation as long as the unitary matrix $\mathbf{V}$ is unknown. The conclusion naturally holds for any related method, such as partial coherences, Cholesky factorization of the CSM, etc. As mentioned in the introduction, this result is not new; yet, as demonstrated in the next section, its derivation in the domain of the source coefficients (rather than in the domain of the measured pressures $\mathbf{p}$) will reveal special configurations under which statistical orthogonality of the sources happens to be a sufficient condition for their separation.

3. When are virtual sources coinciding with true sources?

3.1. Problem formulation

Subsection 2.2 discussed the difference between the virtual sources – as obtained from statistical orthogonalization – and the true sources. The question now arises whether there exist physical configurations where the two concepts can coincide. First, the requirement has to be slightly relaxed since in most practical instances one will content oneself with equality of the partial *powers* of the virtual and true sources (rather than the sources themselves). Formally, the objective is to arrive at

$$\pi_i^{virtual}(\mathbf{r}) = \pi_i(\mathbf{r}), i = 1, \ldots, N_s. \tag{20}$$

Thus, the virtual sources are allowed to be different from the true sources up to an arbitrary (but constant) phase φ,

$$s_i^{virtual}(\mathbf{r}) = s_i(\mathbf{r})e^{i\varphi}, i = 1, \ldots, N_s, \tag{21}$$

which is reminiscent to the classical indeterminacies of BSS [54]. From the results of the previous section (see Eq. (19)) it is clear that such a configuration imposes the constraint $\mathbf{V} = \boldsymbol{\Pi}$ with $\boldsymbol{\Pi}$ a diagonal matrix with unit-magnitude complex entries.

It is now proved that the scenario corresponds to the case where the sound sources are spatially orthogonal.

3.2. A sufficient condition: spatially orthogonal sources

Let us consider the configuration where the SOIs are spatially orthogonal, that is when the following scalar products of any pair of two distinct sources is nil,

$$\langle s_i, s_j \rangle_{\mathbf{M}} = \int_{\Gamma} s_i^*(\mathbf{r}) s_j(\mathbf{r}) M(\mathbf{r}) d\Gamma(\mathbf{r}) = 0, \forall i \neq j \tag{22}$$

with $M(\mathbf{r})$ a given metric introduced for the sake of generality. Resorting to the discretization scheme introduced in Eq. (4), this is equivalently expressed as

$$\langle s_i, s_j \rangle_{\mathbf{M}} = \alpha_i \alpha_j \mathbf{e}_i^T \mathbf{\Psi}^H \mathbf{M} \mathbf{\Psi} \mathbf{e}_j = 0, \quad \forall i \neq j \tag{23}$$

in terms of the matrix of mode shapes,

$$\mathbf{\Psi} = \begin{bmatrix} \varphi_1(r_1) & \cdots & \varphi_{N_s}(r_1) \\ \vdots & & \vdots \\ \varphi_1(r_N) & \cdots & \varphi_{N_s}(r_N) \end{bmatrix}, \tag{24}$$

and

$$\mathbf{M} = \begin{bmatrix} M(r_1)\Delta\Gamma(r_1) & & \\ & \ddots & \\ & & M(r_N)\Delta\Gamma(r_N) \end{bmatrix}, \tag{25}$$

a diagonal "metric matrix". Typical choices for the metric are the identity, $\mathbf{M} = \mathbf{I}$, or the Gramian of the propagation operator, $\mathbf{M} = \mathbf{G}^H\mathbf{G}$, in which case spatial orthogonality is not assessed in the source domain but in an arbitrary domain to which sources are radiated (e.g. the measurement domain). Eq. (23) states the spatial orthogonality of the sources is equivalent to having their mode shapes orthogonal with respect to a given metric. As to be seen in Fig. 2(b), a simple case of this instance is when the sources are disjoint in space – a case of special interest which will be considered in the next section; another case is when the sources have disjoint spatial spectra (i.e. when they are disjoint in the wavenumber domain). In order to investigate the effect of spatial orthogonality on virtual sources, let us now express Eq. (23) in terms of the spatial basis $\mathbf{\Phi}$,

$$\alpha_i \alpha_j \mathbf{e}_i^T \mathbf{A}^H \mathbf{\Phi}^H \mathbf{M} \mathbf{\Phi} \mathbf{A} \mathbf{e}_j = 0, \quad \forall i \neq j. \tag{26}$$

Since the above equation must be true whatever the values of the latent variables α_i and α_j, it implies that $\mathbf{A}^H\mathbf{\Phi}^H\mathbf{M}\mathbf{\Phi}\mathbf{A}$ is a diagonal matrix,

$$\mathbf{A}^H \mathbf{\Phi}^H \mathbf{M} \mathbf{\Phi} \mathbf{A} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{N_s} \end{bmatrix} = \mathbf{\Lambda}. \tag{27}$$

Finally, substituting $\mathbf{A}$ for its SVD given in Eq. (13), the above equation becomes

$$\mathbf{V}\left(\mathbf{D}\mathbf{U}^H \mathbf{\Phi}^H \mathbf{M} \mathbf{\Phi} \mathbf{U} \mathbf{D}\right) \mathbf{V}^H = \mathbf{\Lambda}. \tag{28}$$

This is the main result of this section. It states that spatial orthogonality of the sources is equivalent to having the unitary matrix $\mathbf{V}$ diagonalizing the product of matrices $\mathbf{D}\mathbf{U}^H\mathbf{\Phi}^H\mathbf{M}\mathbf{\Phi}\mathbf{U}\mathbf{D}$. This can be exploited in two ways. First, it provides a recipe to estimate $\mathbf{V}$, which will be further investigated in Section 4. Second, it returns a sufficient condition for virtual sources to be equal to the true sources. As mentioned above this condition requires $\mathbf{V} = \mathbf{\Pi}$ which, upon insertion into Eq. (28), gives

$$\mathbf{D}\mathbf{U}^H \mathbf{\Phi}^H \mathbf{M} \mathbf{\Phi} \mathbf{U} \mathbf{D} = \underbrace{\mathbf{\Pi}^{-1} \mathbf{\Lambda} \mathbf{\Pi}^{-H}}_{\text{diagonal}}. \tag{29}$$

Equality (29) can be satisfied in several ways. i) A simple configuration is when the metric is proportional to the identity, $\mathbf{M} \propto \mathbf{I}$, and the spatial basis is canonical, $\mathbf{\Phi} = \mathbf{I}$, such as with beamforming, or orthonormal, $\mathbf{\Phi}^H\mathbf{\Phi} = \mathbf{I}$, such as in classical NAH. In both cases $\mathbf{\Phi}^H\mathbf{M}\mathbf{\Phi} \propto \mathbf{I}$ and therefore $\mathbf{D}\mathbf{U}^H\mathbf{\Phi}^H\mathbf{M}\mathbf{\Phi}\mathbf{U}\mathbf{D} \propto \mathbf{D}^2$ is diagonal. ii) Another configuration is with the optimal spatial basis introduced in Ref. [9] which, by construction, satisfies $\mathbf{\Phi}^H\mathbf{\Sigma}^{-1}\mathbf{\Phi} = \mathbf{I}$ with $\mathbf{\Sigma}$ a so-called "aperture function". iii) If it happens that the actual mode shapes of the SOIs are known, then orthogonality is satisfied by setting $\mathbf{\Phi} = \mathbf{\Psi}$. iv) Finally, a more general case is when $\mathbf{\Phi}^H\mathbf{M}\mathbf{\Phi} = \mathbf{U}\mathbf{\Delta}\mathbf{U}^H$ where $\mathbf{\Delta}$ stands for a diagonal matrix, which has less obvious physical meaning.

To summarize and keep only the physically most relevant configurations, the virtual sources equal the true sources (in the sense of Eq. (21)) if

1) the sources are spatially orthogonal in a given domain as specified by Eq. (22),
2) and the spatial basis $\mathbf{\Phi}$ is chosen orthonormal in that domain (i.e. $\mathbf{\Phi}^H\mathbf{M}\mathbf{\Phi} \propto \mathbf{I}$).

It is highlighted that the two above conditions must hold simultaneously. Although condition (2) is less natural than condition (1), it reflects the fact that the eigen-elements **U** and **D** returned by the EVD in Eq. (14) generally depend on the choice of the spatial basis $\mathbf{\Phi}$. In other words, the results of PCA depend on the domain in which they are obtained.

Noteworthy is that condition (2) holds generally, in particular not only when PCA is applied on the source coefficients **c**, but also when it is applied directly on the measured pressures **p**, as seen when setting $\mathbf{M}=\mathbf{G}^H\mathbf{G}$. It is believed that this result explains some of the successes of PCA for BSS that have been reported in the literature, for instance when using the canonical basis $\mathbf{\Phi}=\mathbf{I}$ and in the far field where $\mathbf{G}^H\mathbf{G}$ tends to diagonality.

A particular case of spatial orthogonality of the SOIs that has a high physical relevance is reported in the next subsection.

3.3. A particular case: spatially disjoint sources

The configuration with spatially disjoint sources (i.e. sources with non-overlapping support sets in physical space) is a particular case of spatially orthogonal sources with practical interest. Indeed, there are many instances where sound sources are likely to originate from different (disjoint) spatial regions, even though their radiated fields rapidly overlap in space at some distance from the source domain. In mathematical terms, Eq. (22) then naturally holds for any metric **M** since the product of the mode shapes of disjoint sources is always zero. This configuration is investigated in some depth in the next section in order to devise a simple BSS algorithm. Following similar lines, an algorithm could also be developed to separate sources with disjoint spatial spectra, which is another instance of spatial orthogonality, yet not investigated in this paper.

4. Blind separation of spatially disjoint sources

4.1. 4.1. Objectives

The previous section has established a sufficient condition under which virtual sources equal the SOIs. The result is particularly interesting for separating disjoint sources, which is a likely instance of spatially orthogonal sources. In theory, PCA is therefore enough to solve the BSS problem. However, in practice, the condition of perfectly disjoint supports may be difficult to reach for different reasons. First, backpropagation is never perfect and always involve some spatial leakage. Second, the exact topography of the source domain is hardly known in practice and the sources are backpropagated on a surface close to it but not coinciding with it. This necessarily involves radiation from one SOI towards the other SOIs (see the intermediate plane in Fig. 1). Third, the notion of a compact support set may be only approximate in reality, such as with aeroacoustic sources. A last but not least limitation is the requirement to use an orthonormal spatial basis $\mathbf{\Phi}$.

For all these reasons, PCA alone may not be the most effective method when trying to separate spatially disjoint sources. Based on the results of the previous section, a more robust BSS method is to recover SOIs which are forced to be statistically orthogonal (by using PCA) and, at the same time, as disjoint as possible. This will be referred to as joint statistical and spatial orthogonalization in the following.

4.2. Enforcing statistical and spatial orthogonality

In this section, an algorithm is designed to blindly separate incoherent and disjoint sound source independently of the choice of the basis functions $\mathbf{\Phi}$ used in the backpropagation step.

The starting point of the proposed algorithm is Eq. (28) where it is seen that the unknown matrix **V** that relates the virtual sources to the true ones (see Eq. (19)) actually diagonalizes the product of matrices $\mathbf{DU}^H\mathbf{\Phi}^H\mathbf{M\Phi UD}$ wherein all quantities are known. For disjoint sources, this is true whatever the metric **M**. Thus, in theory, **V** can be uniquely recovered. Going back to Eq. (23), this enforces the diagonality of the matrix product $\mathbf{\Psi}^H\mathbf{M\Psi}$ for any metric **M** and therefore the spatial orthogonality (22). This novel criterion is referred to as "spatial orthogonalization". The idea is illustrated on one-dimensional signals in Fig. 4.

Therefore, the proposed algorithm is the following:

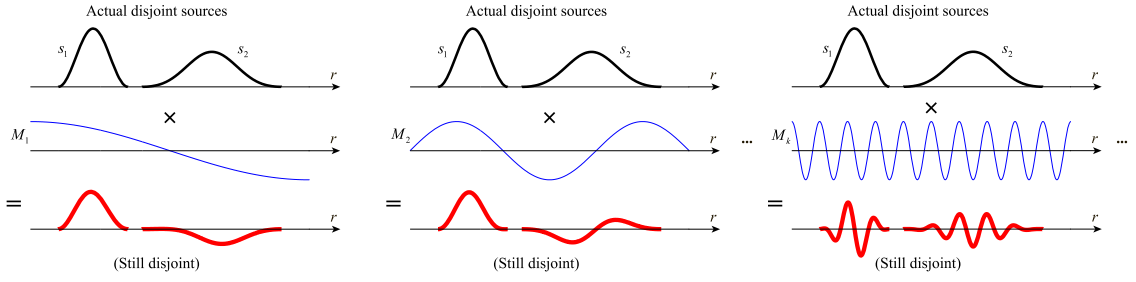

Fig. 4. Principle of the criterion of spatial orthogonalization: when disjoint, two source distributions $s_1(\mathbf{r})$ and $s_2(\mathbf{r})$ always have zero scalar product even if pre-multiplied by an arbitrary spatial function $M_k(\mathbf{r})$. This is illustrated here with $M_k(\mathbf{r})$ generated by sinusoids with increasing wavenumbers.

1) Statistical orthogonalization: estimate the eigen-elements **U** and **D** from the EVD (i.e. PCA) of the CSM $\mathbb{E}_{\varpi}\{\hat{\mathbf{c}}\hat{\mathbf{c}}^H\}$ as conducted in Eq. (14),
2) Construct a set of matrices $\left\{\mathbf{D}\mathbf{U}^H\mathbf{\Phi}^H\mathbf{M}_k\mathbf{\Phi}\mathbf{U}\mathbf{D}\right\}_{k=1}^{K}$ for different candidate metrics $\mathbf{M}_k$,
3) Spatial orthogonalization: find the unitary matrix **V** that jointly diagonalizes the set of matrices $\left\{\mathbf{D}\mathbf{U}^H\mathbf{\Phi}^H\mathbf{M}_k\mathbf{\Phi}\mathbf{U}\mathbf{D}\right\}_{k=1}^{K}$,
4) Recover the individual sources $\mathbf{s}_i = \mathbf{\Phi}\mathbf{U}\mathbf{D}\mathbf{V}^H\mathbf{e}_i$.

It is noteworthy that the algorithm can be used with any choice of the spatial basis $\mathbf{\Phi}$ – this includes NAH and beamforming as particular cases, but not only. On the contrary, the metrics $\mathbf{M}_k$ should be carefully chosen. Here, the harmonic functions are recommended, although other orthogonal polynomials such as Legendre and Chebyshev have been tested with success. The definition of the harmonic functions is

$$M_k(\mathbf{r}) = \exp(J\mathbf{r}^T\mathbf{k}_k), \quad J^2 = -1, \tag{30}$$

where $\mathbf{k}_k$ denotes the k-th element of a set of arbitrary wavevectors used to synthesize oscillating functions embedded in the metric $\mathbf{M}_k$ – see the second row in Fig. 4. How to jointly diagonalize the set of matrices in Step (3) of the algorithm is addressed in the following subsection.

4.3. Joint approximate diagonalization algorithm

Let us define

$$C_{ij}^{[k]}(\mathbf{V}) = \mathbf{e}_j^T\mathbf{V}\mathbf{D}\mathbf{U}^H\mathbf{\Phi}^H\mathbf{M}_k\mathbf{\Phi}\mathbf{U}\mathbf{D}\mathbf{V}^H\mathbf{e}_i. \tag{31}$$

The joint diagonalization of the set of weighted spatial correlation coefficients $\left\{\mathbf{C}^{[k]}(\mathbf{V})\right\}_{k=1}^{K}$ may be naturally achieved by minimizing the sum of the squared magnitudes of their off-diagonal elements, that is

$$\hat{\mathbf{V}} = \underset{\mathbf{V}}{\text{Argmin}} \left\{ \sum_{k=1}^{K} \left| \text{Off}\left\{\mathbf{C}^{[k]}(\mathbf{V})\right\} \right|^2 \right\}, s.\text{t.} \mathbf{V}\mathbf{V}^H = \mathbf{I}, \tag{32}$$

where operator "Off" zeroes the diagonal elements of a matrix. Fortunately, the cost function in Eq. (32) is common in BSS and enjoys several optimization algorithms. Here, the Joint Approximate Diagonalization (JAD) algorithm of Ref. [55] has been used.

Similarly to most BSS methods, the proposed one also needs the determination of the number of sources before separation. The authors suggest the use of the methods introduced in Section 3 of Ref. [52], especially the criterion of "entropic L-curve", although other criteria may also be considered.

5. Experimental validation

This section validates the separation of incoherent and disjoint sources on two experimental setups. The first one involves piston-like sources produces by loudspeakers. The second one deals with the more difficult case of aeroacoustic sources generated by an airfoil submitted to a turbulent flow.

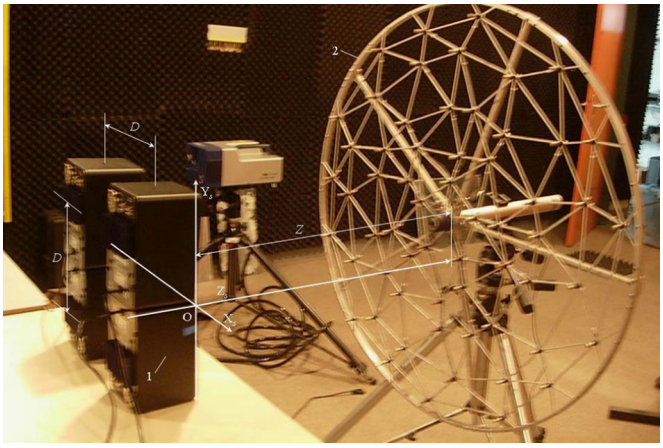

Fig. 5. The configuration of experiments conducted in a semi-anechoic chamber (1 – the loudspeaker; 2 – the slice wheel array of microphones).

5.1. First experimental setup: piston-like sources

The experiments were conducted in a semi-anechoic chamber. Their configuration is depicted in Fig. 5. Four loudspeakers produced incoherent and wideband random sound sources with frequency spectra ranging up to 8 kHz. Two range distances of 10 cm and 100 cm between the source domain of interest and the slice wheel array of microphones were used. These two distances will be referred to as near-field and far-field, respectively, where near-field is understood here as the area that is closer to the sound source than one wavelength (this corresponds to the highest frequencies of 3400 Hz and 340 Hz, respectively). The spacing D between adjacent loudspeakers is 12 cm in the near-field and 18 cm in the far-field. The sampling frequency was set to 16.384 kHz and recording time to 4 min. Measurements were also performed with only a single loudspeaker switched on in order to obtain references for the true sources and allow comparisons. More details on the experimental apparatus can be found in Subsection 4.1 of Ref. [52].

Data recorded in the near-field and in the far-field were then processed by the two proposed approaches: Statistical Orthogonality Only (SO) by PCA and joint Statistical and Spatial Orthogonality (2SO) by the algorithm proposed in Subsection 4.2 with $(N_s+1)^2$ harmonics functions generated with increasing wavenumbers – see Eq. (30). In all cases, the backpropagation method described in Ref. [9] was used. First, the separation results are presented in the whole working frequency band by means of the quadratic velocity spectrum (i.e. Eq. (9)). Second, the quality of the spatial distribution of the separated sources is assessed by means of the spatial correlation spectrum between the separated, $\hat{s}_i$, and the true sources, s_i, as a function of frequency,

$$\rho_i(\omega) = \frac{\sum_{l=1}^{N} \mathbb{E}_{\varpi}\left|\hat{S}_i(\mathbf{r}_l,\omega)S_i^*(\mathbf{r}_l,\omega)\right|}{\sqrt{\sum_{l=1}^{N}\mathbb{E}_{\varpi}\left|\hat{S}_i(\mathbf{r}_l,\omega)\right|^2 \sum_{l=1}^{N}\mathbb{E}_{\varpi}\left|S_i^*(\mathbf{r}_l,\omega)\right|^2}}. \tag{33}$$

5.2. Separation results from statistical orthogonalization

As demonstrated in Section 3.2, forcing only the statistical orthogonalization of the sources by diagonalizing the CSM may return satisfactory separation results provided that the sound sources have spatially orthogonal distributions. This is experimentally verified in the next subsection.

5.2.1. Source separation in the near-field

In the near-field, the working frequency is set to 1800 Hz (as will be seen shortly, this corresponds to a frequency range where SO works well; it also corresponds to the frequency used in Ref. [52] so as to allow comparisons). The separation

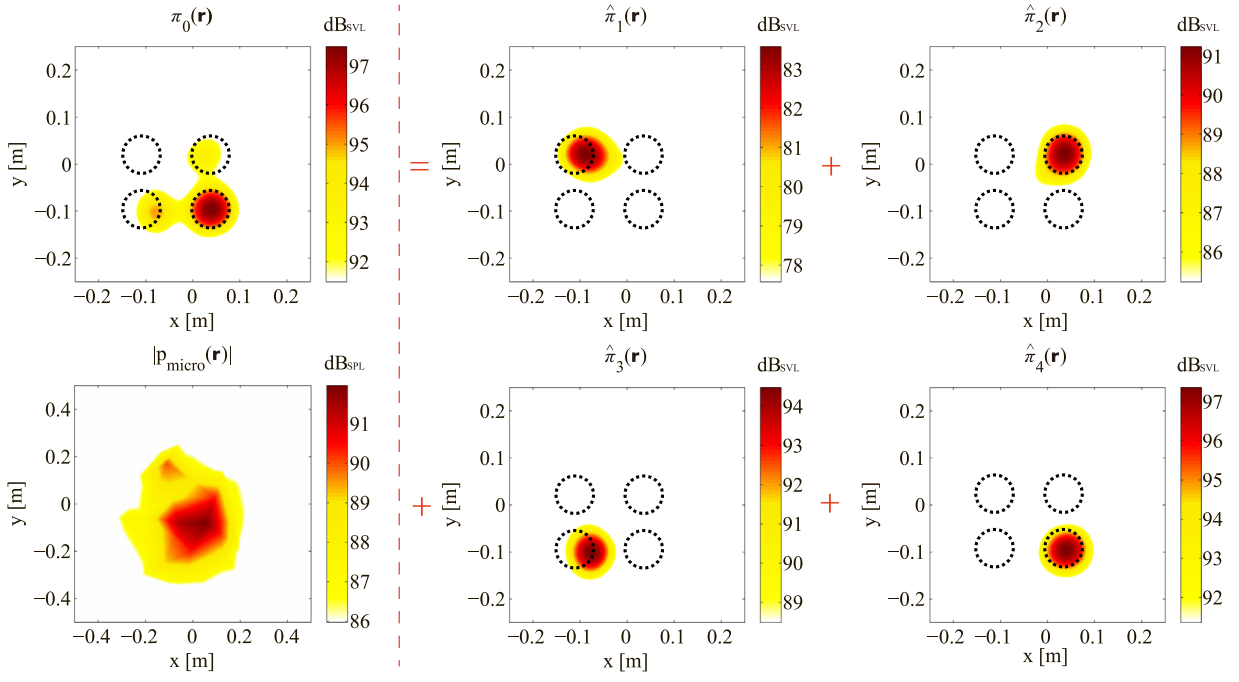

Fig. 6. Separated sources from statistical orthogonalization in the near-field at 1800 Hz. The total quadratic velocity $\pi_0(\mathbf{r},\omega)$ is decomposed into components $\hat{\pi}_i(\mathbf{r})$, $i = 1,\ldots,4$. The dynamic range is held constant to 6 dB whereas scale is adapted to allow the visualization of the smallest sources.

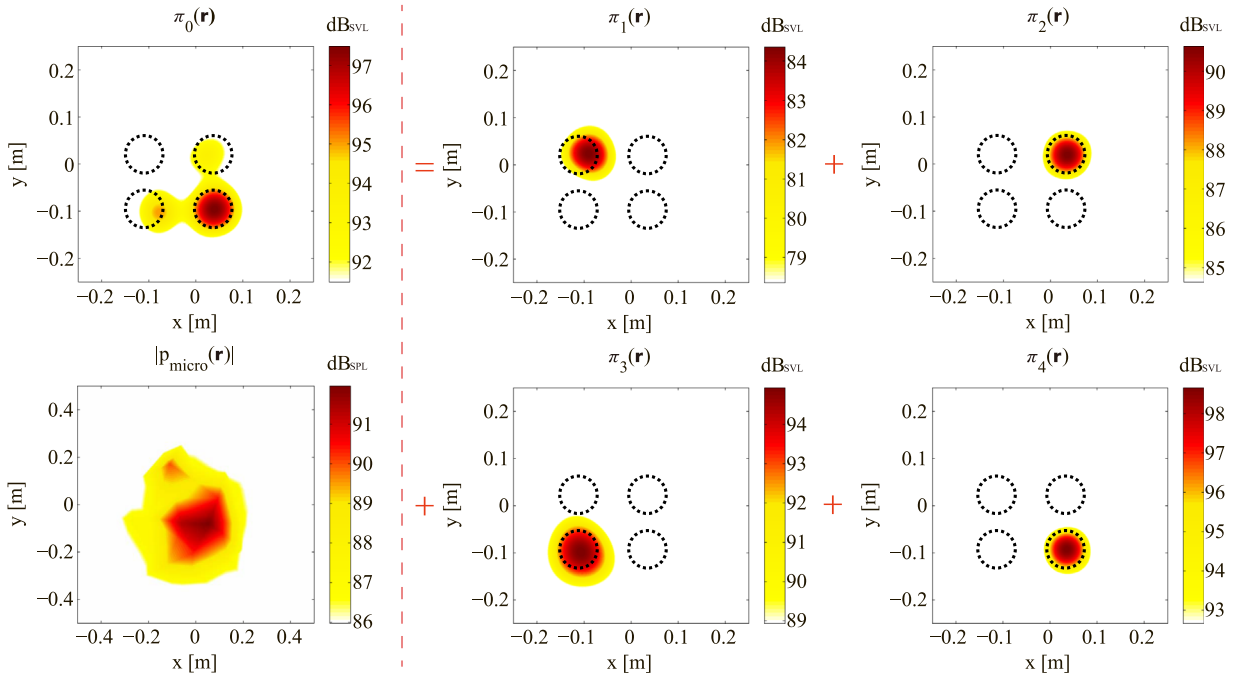

Fig. 7. True sources obtained from individual measurements in the near-field at 1800 Hz. (Dynamic range = 6 dB).

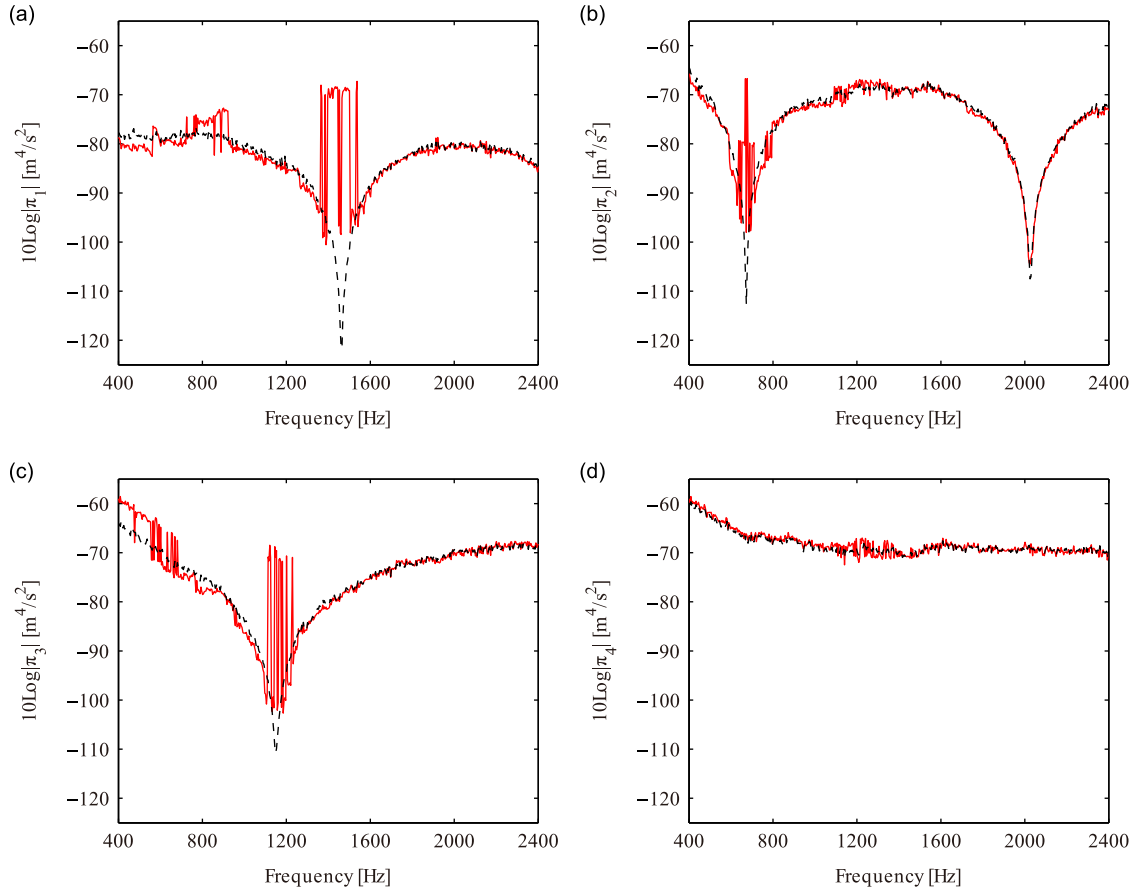

Fig. 8. Quadratic velocity spectra of the separated sources, (a) $\hat{\pi}_1(\omega)$, (b) $\hat{\pi}_2(\omega)$, (c) $\hat{\pi}_3(\omega)$, and (d) $\hat{\pi}_4(\omega)$, from simple statistical orthogonalization (red solid line) compared to true sources (black dashed line) in the near-field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

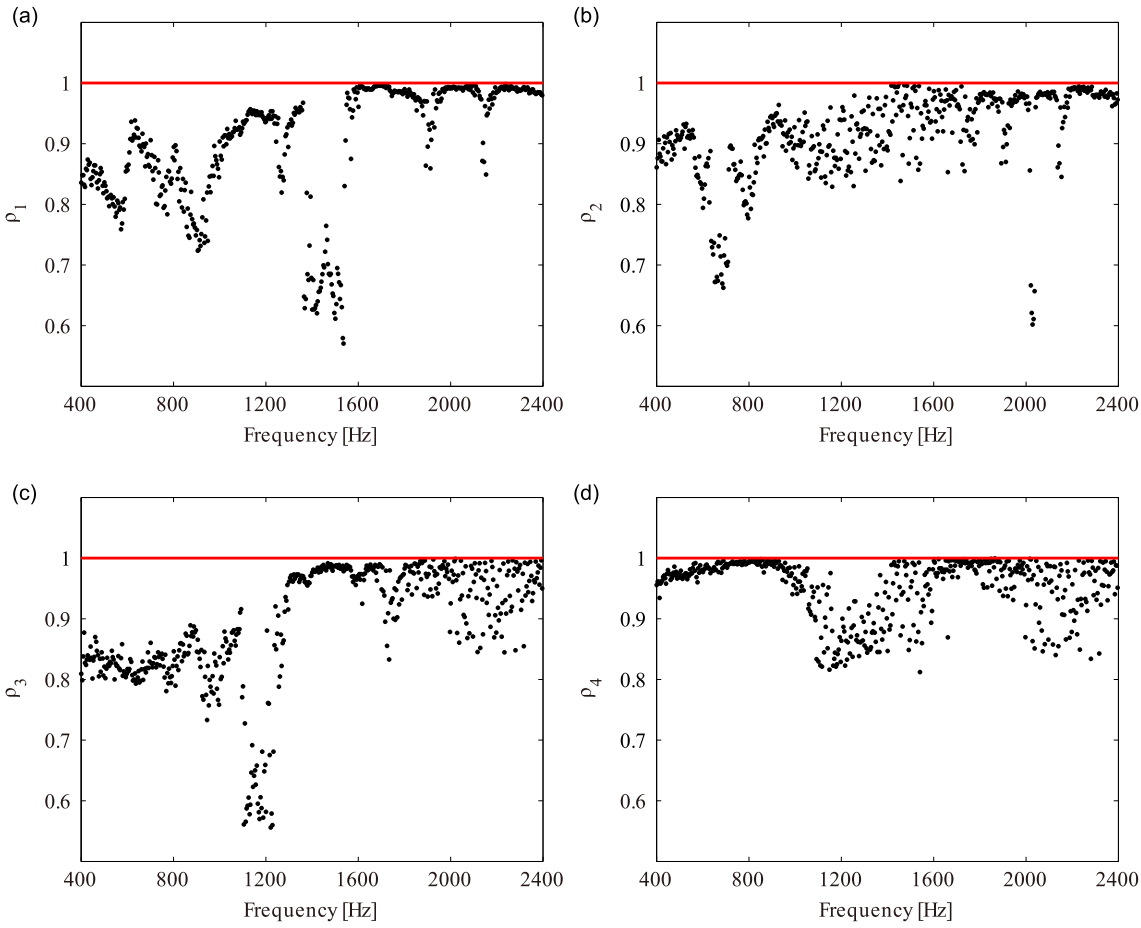

Fig. 9. Spatial correlation spectra between the separated and true sources, (a) $\rho_1(\omega)$, (b) $\rho_2(\omega)$, (c) $\rho_3(\omega)$, and (d) $\rho_4(\omega)$, in the near-field.

results from SO are displayed in Fig. 6 which shows how the quadratic velocity (normal component to the source plane) $\pi_0(\mathbf{r},\omega)$ before separation is decomposed into components $\hat{\pi}_i(\mathbf{r},\omega)$, $i = 1, \ldots, 4$, relating to each source. It appears that all the estimated sources account properly for the sound power and that they could be successfully separated because, at this frequency, their supports are mainly disjoint, a sufficient condition for SO. To evaluate the separation performance, the corresponding four true sources obtained from individual measurements are displayed in Fig. 7. Excellent matching is observed between the separated and the true sources with respect to both magnitude and location estimation.

Next, the separation results are displayed in terms of quadratic velocity spectra in Fig. 8 and spatial correlation spectra in Fig. 9 in the whole available frequency band. In order to solve the permutation ambiguity inherent to BSS (arbitrary labeling of the separated sources at each frequency), the separated sources were re-ordered at each frequency according to the spatial correlation between the separated sources and the true ones. The lower bound of the spectra is determined by the size of the microphone array, whereas the corresponding upper bound depends on the minimum spacing between two adjacent microphones. From Fig. 8 it can be seen that the spectra of the separated sources are consistent with those of the true sources except in the very low frequency interval [400,600] Hz and in the frequency band [1300,1500] Hz which both enclose zeros of the quadratic velocity spectrum of the true source s_1. However, as shown in the spatial correlation spectra (see Fig. 9), there exists some discrepancies even though the power distributions seemed to match well. The highest spatial correlation – above 0.8 – is mostly at high frequencies (above 1600 Hz) where the source supports become more and more disjoint due to finer and finer spatial resolution. The spatial correlation coefficients between the separated sources and the four true ones are 0.99, 0.98, 0.98 and 1 at 1800 Hz.

5.2.2. Source separation in the far-field

The separated and true sources in the far-field are displayed in Figs. 10 and 11, respectively. Separation results are very satisfactory in terms of estimated magnitudes, locations, and spatial distributions. This demonstrates that statistical orthogonalization can actually achieve good spatial separation provided the sources are spatially orthogonal, a condition which is naturally satisfied here when their supports are disjoint.

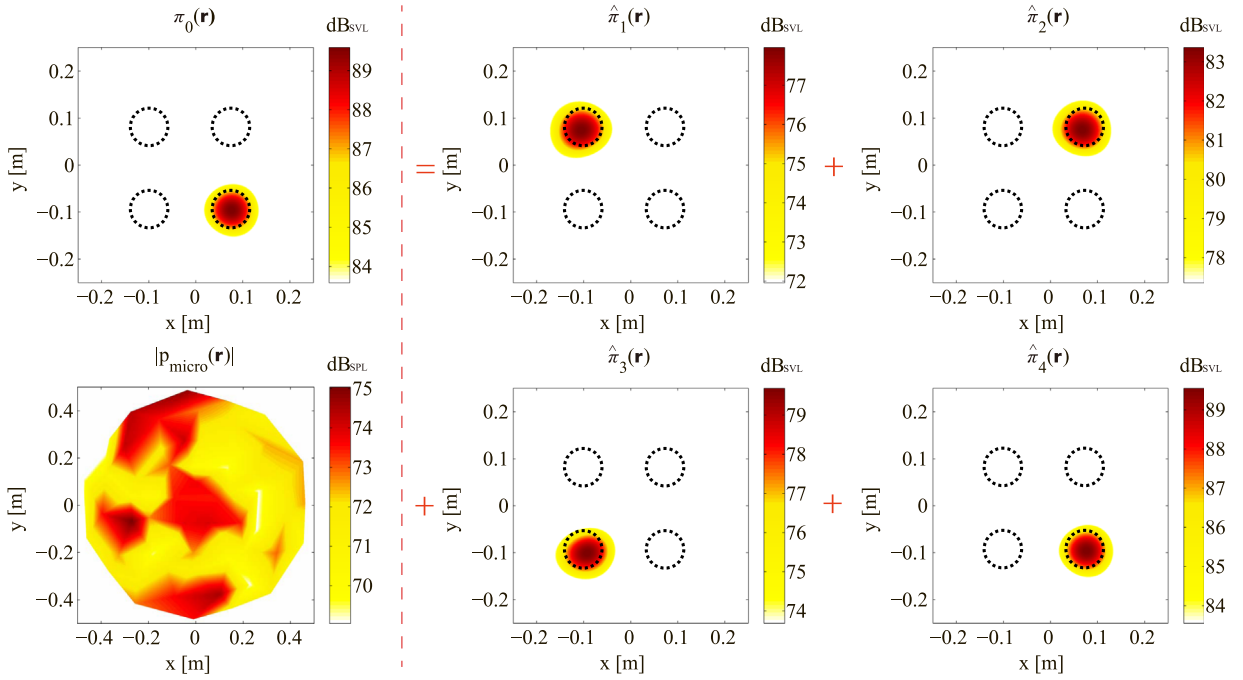

Fig. 10. Separated sources from statistical orthogonalization in the far-field at 3600 Hz. The total quadratic velocity $\pi_0(\mathbf{r}, \omega)$ is decomposed into components $\hat{\pi}_i(\mathbf{r})$, $i = 1,\ldots,4$. (Dynamic range = 6 dB).

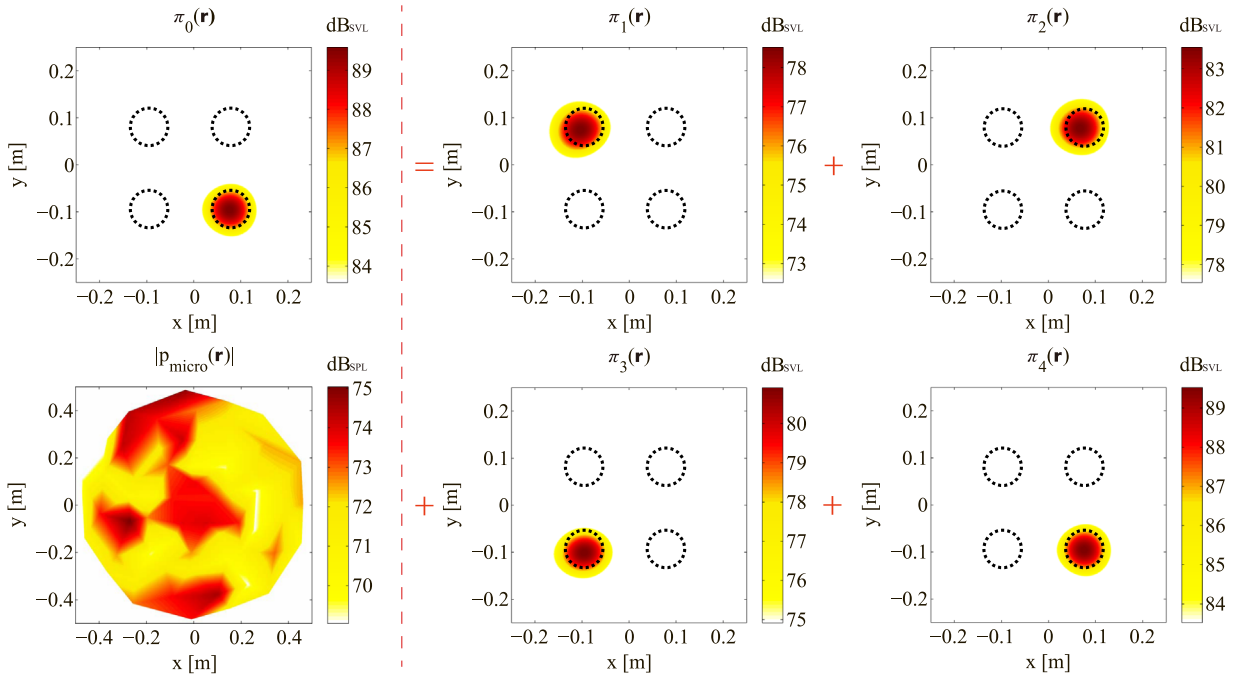

Fig. 11. True sources obtained from individual measurements in the far-field at 3600 Hz. (Dynamic range = 6 dB).

Figs. 12 and 13 display the quadratic velocity spectra and the spatial correlation spectra in the whole available frequency band, respectively. Similarly to the near-field, the boundaries of the whole working frequency band are determined by the characteristic parameters of the microphone array and by the distance Z. The lower working frequency limit depends on the ratio of the distance Z and the size of the applied array (e.g. the diameter of the slice wheel array here). On the contrary, the upper limit of working frequency is inferred by the ratio between the distance Z and the minimum spacing between two adjacent microphones.

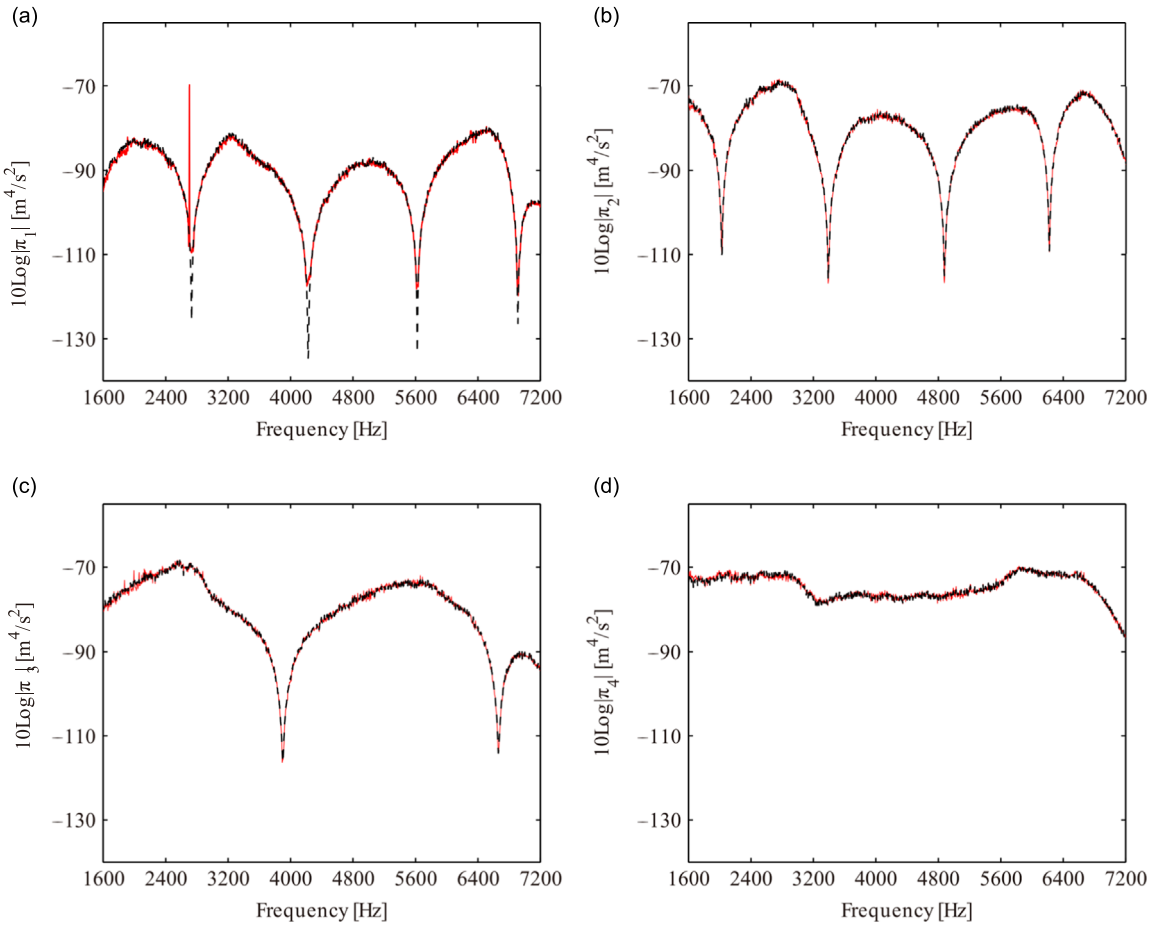

Fig. 12. Quadratic velocity spectra of the separated sources, (a) $\hat{\pi}_1(\omega)$, (b) $\hat{\pi}_2(\omega)$, (c) $\hat{\pi}_3(\omega)$, and (d) $\hat{\pi}_4(\omega)$, from statistical orthogonalization (red solid line) compared to true sources (black dashed line) in the far-field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The quadratic velocity spectra of the separated sources match remarkably well with those of the true sources, even at the zeros of the first source (see Fig. 12(a)). This is due to the larger source spacing – $D = 18$ cm set in the far-field. Fig. 13 shows that, in the entire frequency band, the spatial correlation is most above 0.8. The corresponding spatial correlation coefficients of the four separated sources are 0.98, 1, 0.96 and 1 at 3600 Hz.

5.3. *Separation results from joint statistical and spatial orthogonalization*

When sound sources are disjoint in space, they might be better separated by enforcing *both* statistical and spatial orthogonalization (see Section 4). The experimental parameters of this subsection are the same as that in Section 5.2.

5.3.1. *Source separation in the near-field*

The working frequency is chosen as 833 Hz (all working frequencies have been selected so as to illustrate at best typical cases where statistical orthogonalization alone does not work and needs to be replaced by joint statistical and spatial orthogonalization in Section 5.3; they also correspond whenever possible to the working frequencies used in Ref. [52] so as to allow comparisons). As illustrated in Fig. 14, only one source can be separated by statistical orthogonalization, because, at such a low frequency, the spatial resolution is no longer sufficient to guarantee disjoint supports and thus spatial orthogonality.

To improve separation results, spatial orthogonality is then enforced jointly with the statistical orthogonality as explained in Subsection 4.2. This is illustrated in Fig. 15, which shows an excellent match to the true sources in Fig. 16. There are few errors between the separated and true sources in terms of energy and spatial distribution, although the shapes of the sound sources are slightly distorted.

The example shows that enforcing both statistical and spatial orthogonality is effective, with similar performance as for the method introduced in Ref. [52]. With the same spatial resolution (1 mm) and the same computational power, the former

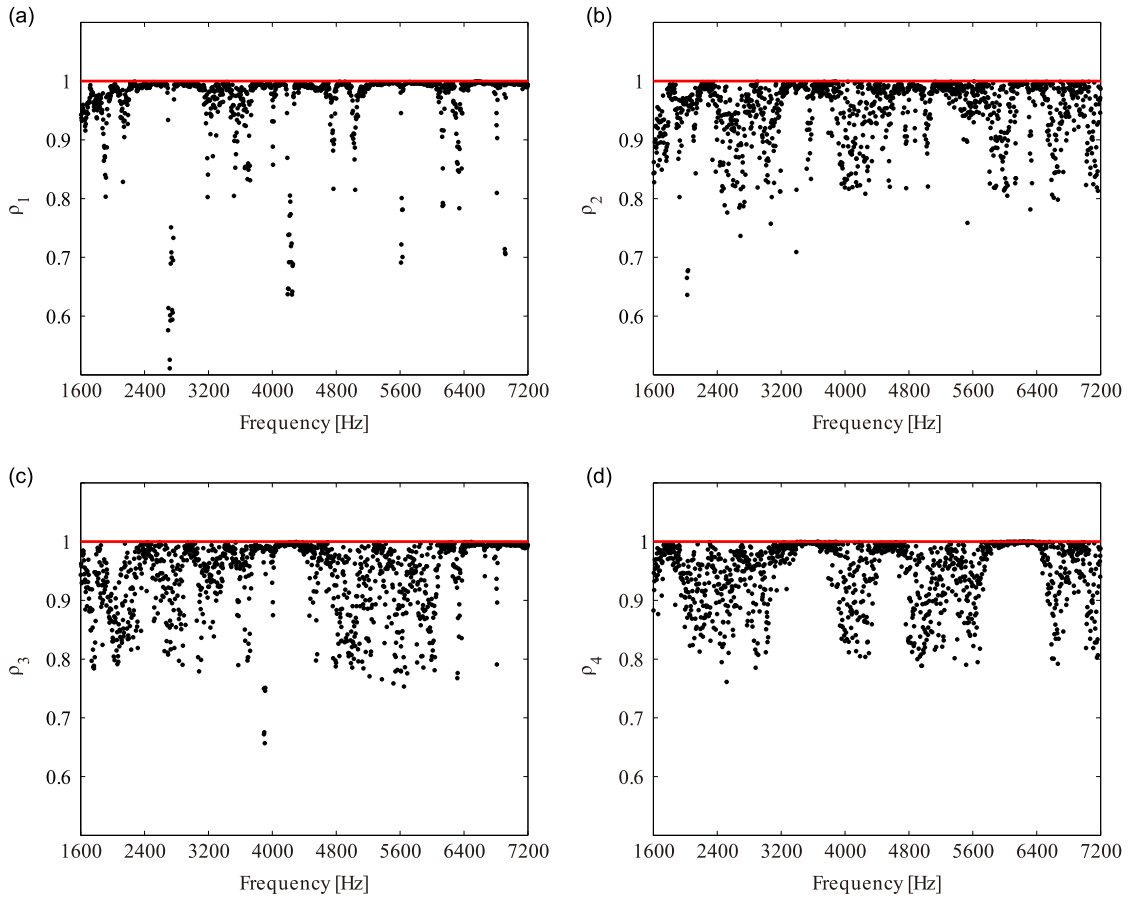

Fig. 13. Spatial correlation spectra between the separated and true sources, (a) $\rho_1(\omega)$, (b) $\rho_2(\omega)$, (c) $\rho_3(\omega)$, and (d) $\rho_4(\omega)$, in the far-field.

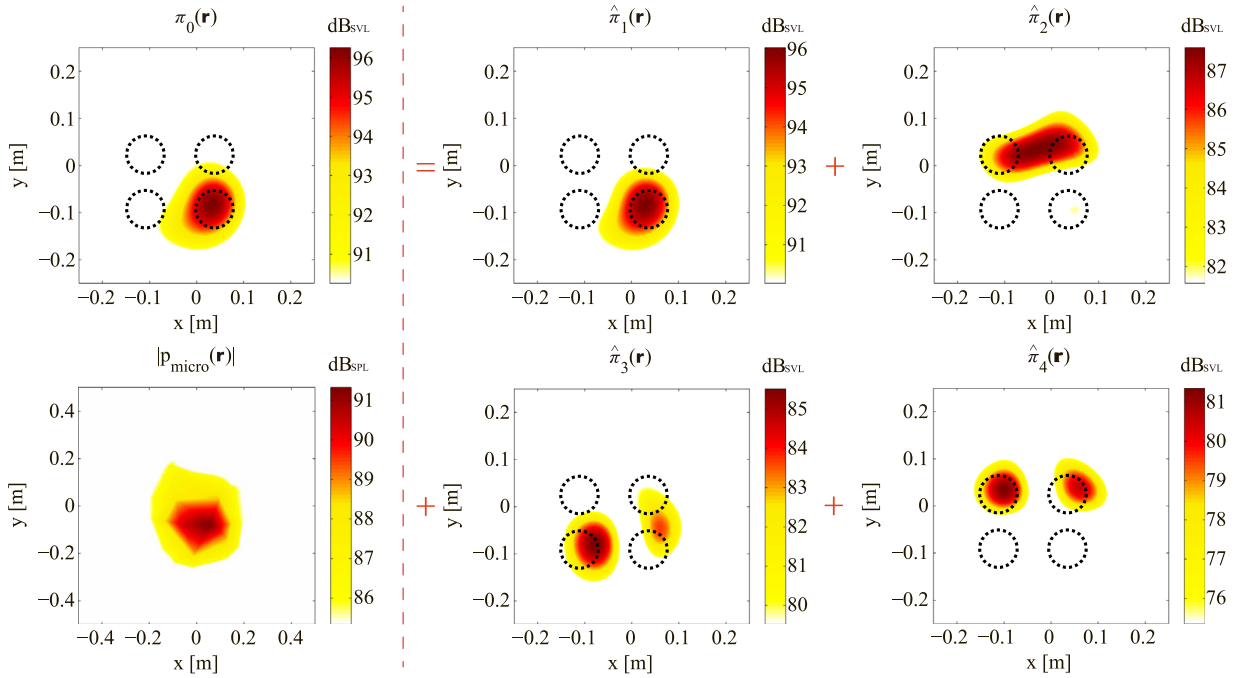

Fig. 14. Separated sources from statistical orthogonalization in the near-field at 833 Hz. The total quadratic velocity $\pi_0(\mathbf{r},\omega)$ is decomposed into components $\hat{\pi}_i(\mathbf{r})$, $i = 1,\ldots,4$. (Dynamic range = 6 dB).

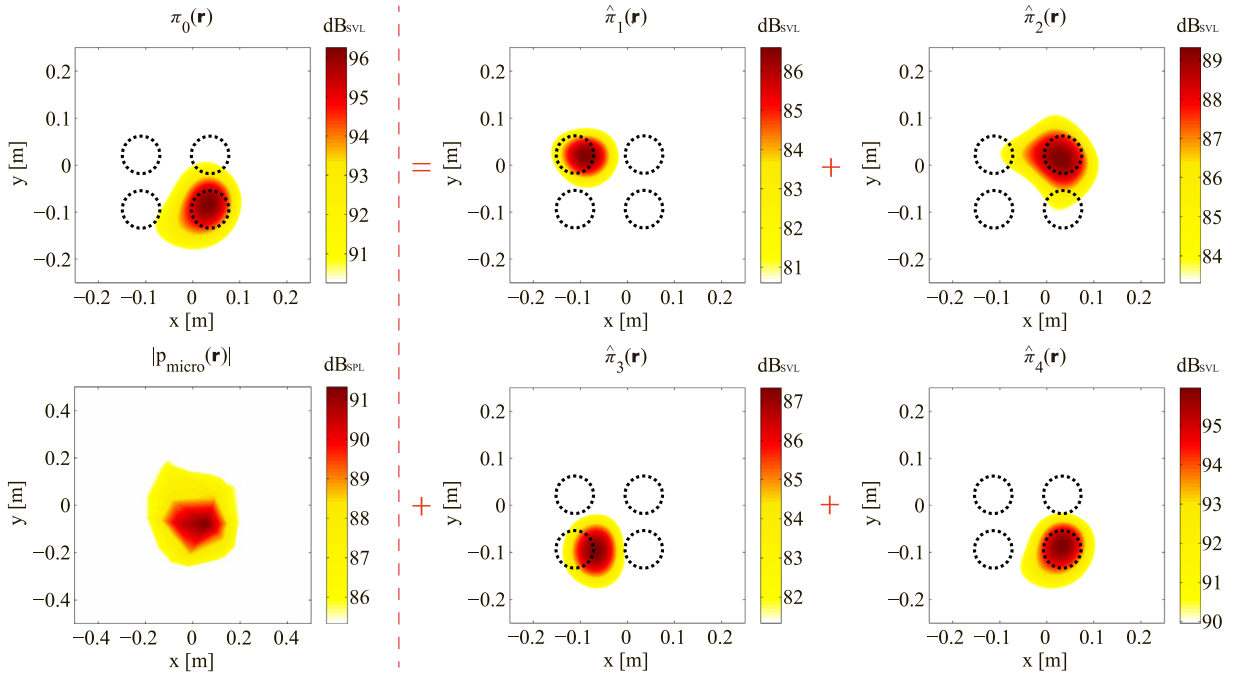

Fig. 15. Separated sources from joint statistical and spatial orthogonalization in the near-field at 833 Hz. The total quadratic velocity $\pi_0(\mathbf{r},\omega)$is decomposed into components $\hat{\pi}_i(\mathbf{r})$, $i = 1,\ldots,4$. (Dynamic range = 6 dB).

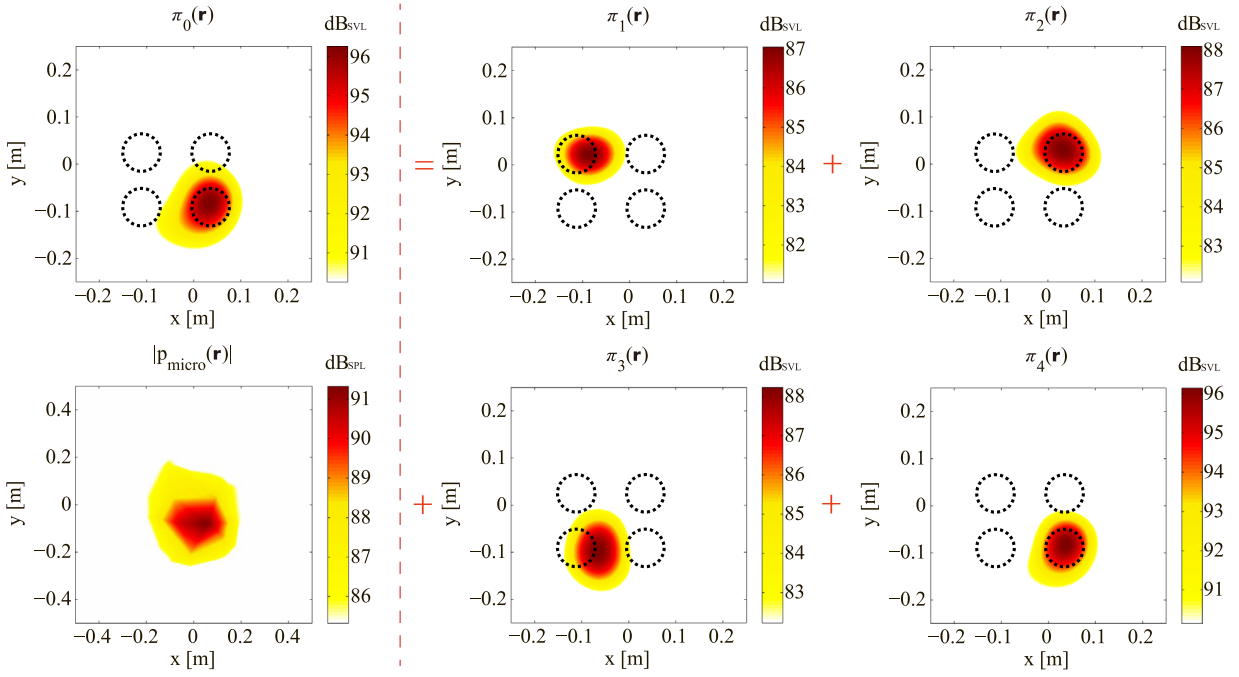

Fig. 16. True sources obtained from individual measurements in the near-field at 833 Hz. (Dynamic range = 6 dB).

achieves virtually identical separation results ten times faster (however it should be kept in mind that the least spatial entropy criterion applies more generally to compact sources, which are not necessarily disjoint).

The quadratic velocity and spatial correlation spectra are shown in Figs. 17 and 18, respectively.

The spatial correlation spectra in Fig. 18 show improved separation sources in the whole working frequency band as compared to simple statistical orthogonalization in Fig. 9. Remarkably, the spatial correlations on the third and fourth true sources are more than 0.98 at most frequencies in the working band (see Fig. 18(c) and (d)). Note however that the spatial

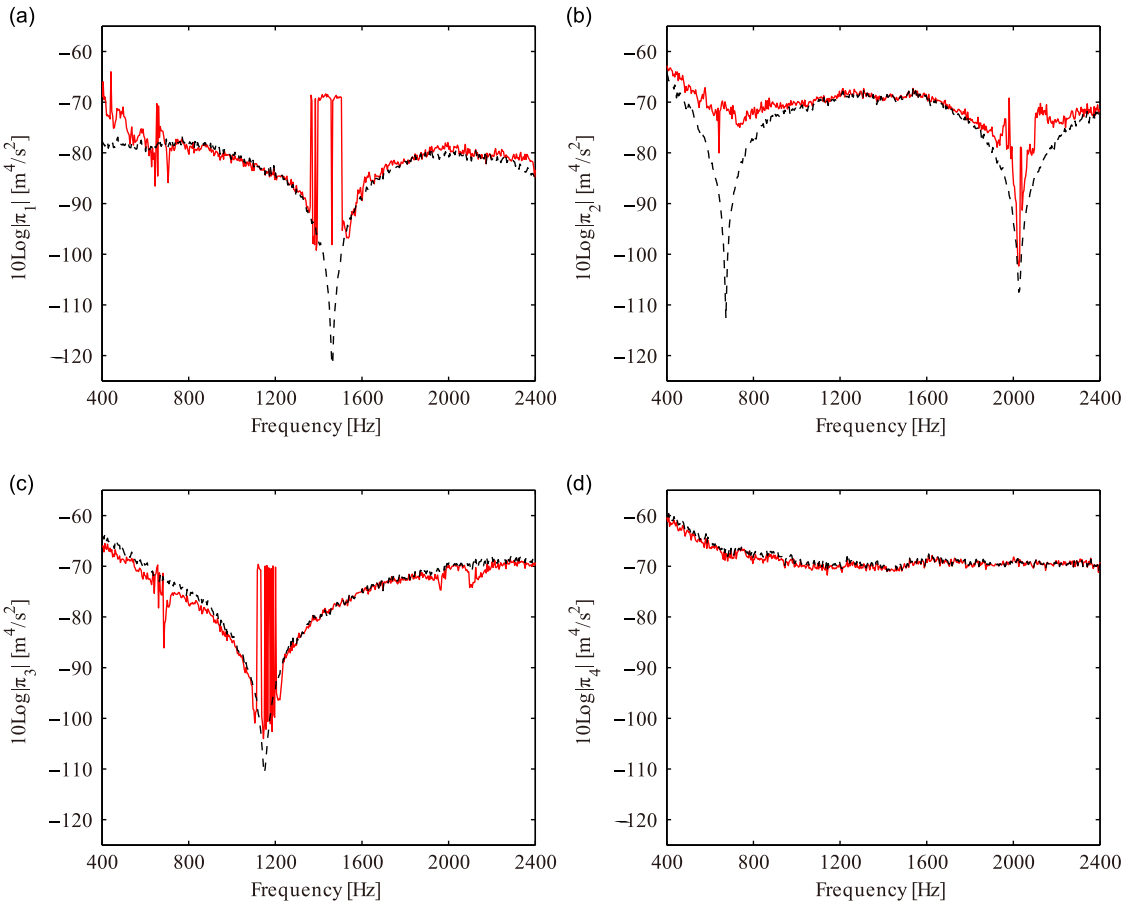

Fig. 17. Quadratic velocity spectra of the separated sources, (a) $\hat{\pi}_1(\omega)$, (b) $\hat{\pi}_2(\omega)$, (c) $\hat{\pi}_3(\omega)$, and (d) $\hat{\pi}_4(\omega)$, from joint statistical and spatial orthogonalization (red solid line) compared to true sources (black dashed line) in the near-field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

correlation drops in some frequency bands where the amplitude of the true source is very small (e.g. around 2 kHz for source 2 in Fig. 18(b)); in such cases, the source is below the noise level and can hardly be separated.

The spatial correlation coefficients between the true and separated sources at 833 Hz are listed in Table 1. It is seen that joint orthogonalization slightly improved the separation results in most cases.

5.3.2. *Source separation in the far-field*

The separated sources from 2SO in the far-field at 2437 Hz are all displayed in Fig. 20. As for the separated sources from simple SO and the true sources, they are presented in Figs. 19 and 21, respectively. The separated sources from 2SO match perfectly the true sources as shown in Fig. 21.

The quadratic velocity and spatial correlation spectra of the separated sources are shown in Figs. 22 and 23. Just as in the near-field case, the quadratic velocity spectra match closely the references except maybe near the zeros and in the low frequency interval [1600,2400] Hz. The good performance of the 2SO separation is also verified by means of the spatial correlation spectra (see Fig. 23).

The spatial correlations coefficients between the separated and true sources at 2437 Hz are listed in Table 2.

5.4. *Second experimental setup: aeroacoustic sources*

In this section the proposed BSS method is employed for the investigation of aeroacoustic sources by means of both simulation and real experiments. In the first part, a numerical experiment is designed to validate the BSS algorithm for aeroacoustic sources. In the second part, results from a real experiment realized in an open-jet anechoic wind tunnel are presented. Both examples aim to illustrate the advantages of blind source separation for the analysis of physical noise sources.

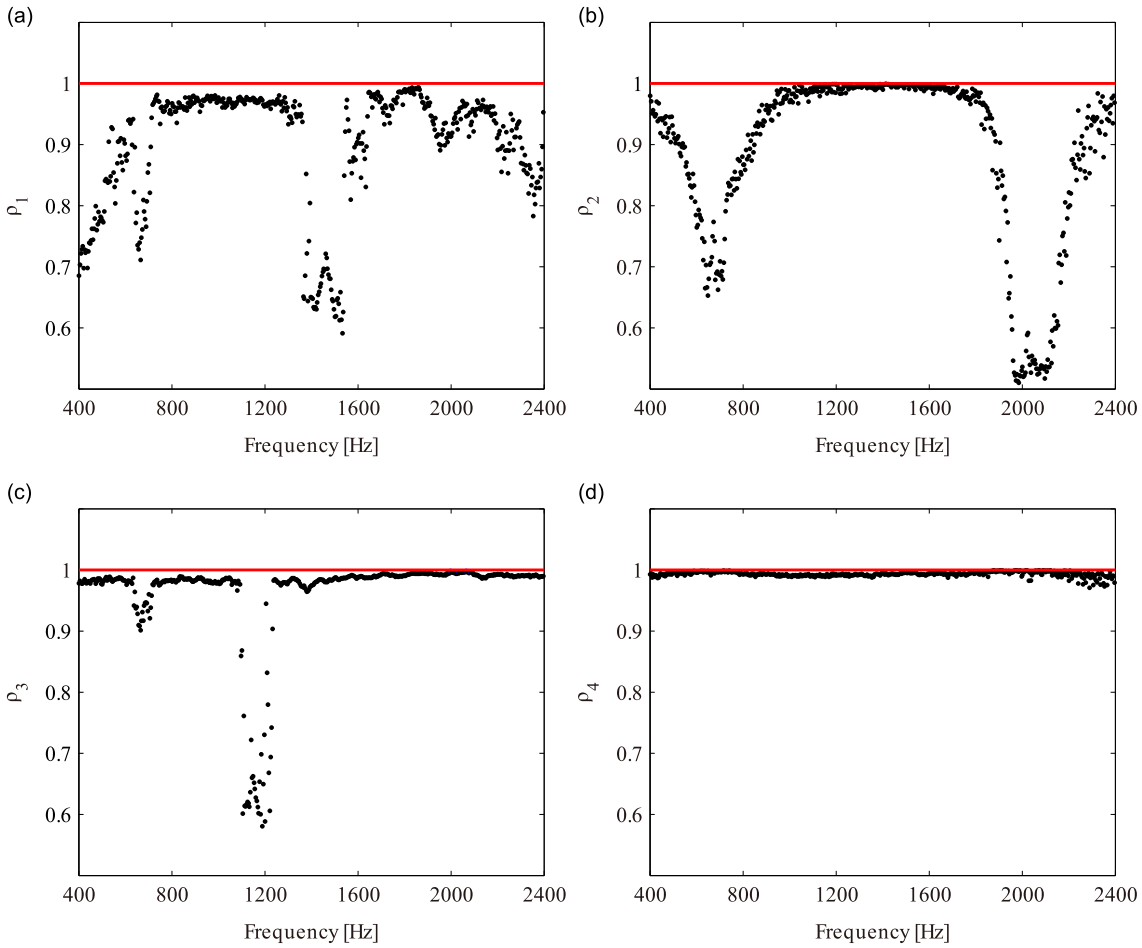

Fig. 18. Spatial correlation spectra between the separated and true sources, (a) $\rho_1(\omega)$, (b) $\rho_2(\omega)$, (c) $\rho_3(\omega)$, and (d) $\rho_4(\omega)$, in the near-field.

Table 1
Spatial correlation coefficients between the separated and true sources from SO and 2SO in the near-field at 833 Hz.

	s_1	s_2	s_3	s_4
SO	0.82	0.88	0.86	1
2SO	0.97	0.91	0.99	0.99

5.4.1. Numerical experiment of spatially distributed sources

The designed simulation is representative of sources encountered in practice in aeroacoustics. That is, spatially distributed sources which may or may not be correlated. The dimensions of the problem (microphone array and source dimensions) match those from the experimental application to be presented in Subsection 5.4.2. The source length L equals the airfoil span ($\sim$30 cm) used in the real experiment. The geometry of the simulation is presented in Fig. 24. In the first case, the distributed source is divided into two groups of mutually correlated sources, each group with a source length equal to $L_s = L/2$. The two groups are mutually uncorrelated. The acoustic pressure is simulated according to the model described in Eq. (4). In this model, the vector $\mathbf{c}$ is a realization (represented here by a snapshot) of a multivariate random process distributed according to $\mathbb{CN}_k(0, \sigma_c \mathbf{\Sigma}_c)$, which is a multivariate complex Gaussian distribution of dimension k, zero mean vector and with a covariance matrix equal to $\sigma_c \mathbf{\Sigma}_c$. The degree of correlation between sources is modeled by the matrix of correlation coefficients $\mathbf{\Sigma}_c$, which is shown in Fig. 25 for the two simulation cases. The same variance σ_c is assigned to all source coefficients. Additive noise is also added to the simulated acoustic pressure, according to Eq. (4). The noise vector $\boldsymbol{\epsilon}$ is drawn from the multivariate complex Gaussian distribution $\mathbb{CN}_M(0, \sigma_n \mathbf{I}_M)$, where σ_n is the variance of the noise and $\mathbf{I}_M$ is the identity matrix of dimension M. It is thus assumed that the noise is uncorrelated between microphones. The number of snapshots is set to 400 and the values of σ_c and σ_n are chosen such as to keep an SNR of 40 dB.

The spatial distributions of sources are reconstructed at the source plane by forcing statistical orthogonality only and by jointly forcing statistical and spatial orthogonality. Results are shown in Figs. 26–28 for different frequencies or, equivalently,

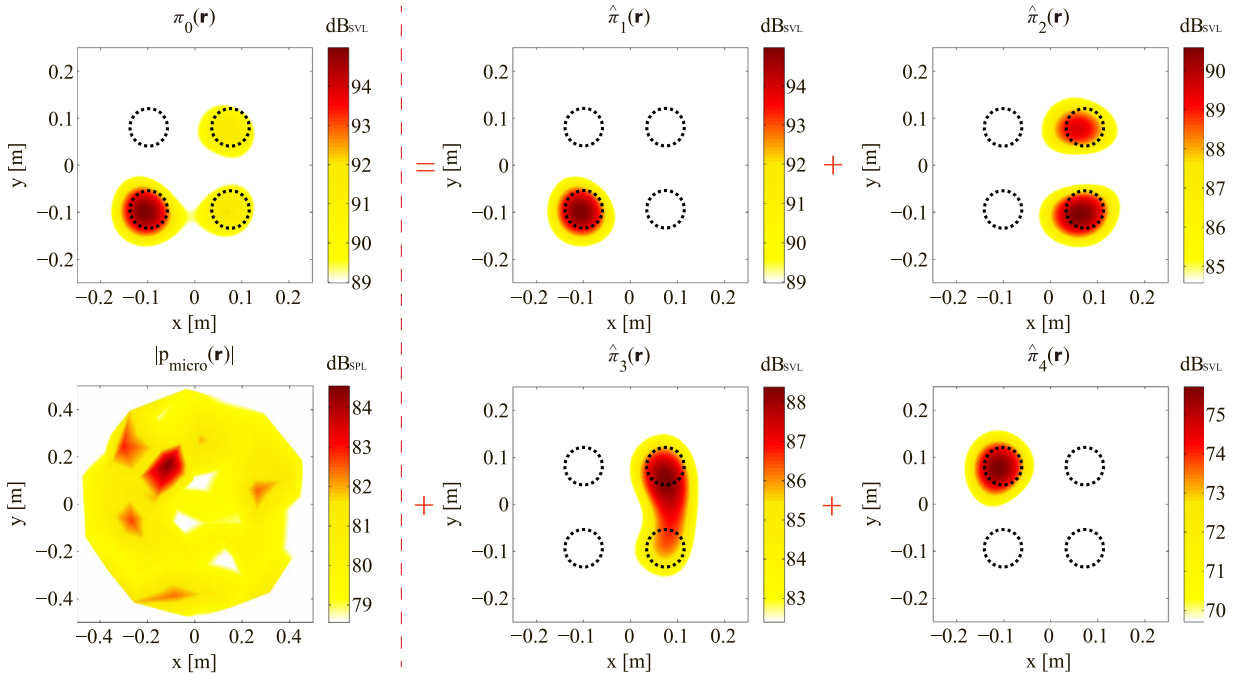

Fig. 19. Separated sources from statistical orthogonalization in the far-field at 2437 Hz. The total quadratic velocity $\pi_0(\mathbf{r},\omega)$ is decomposed into components $\hat{\pi}_i(\mathbf{r})$, $i = 1,\ldots,4$. (Dynamic range = 6 dB).

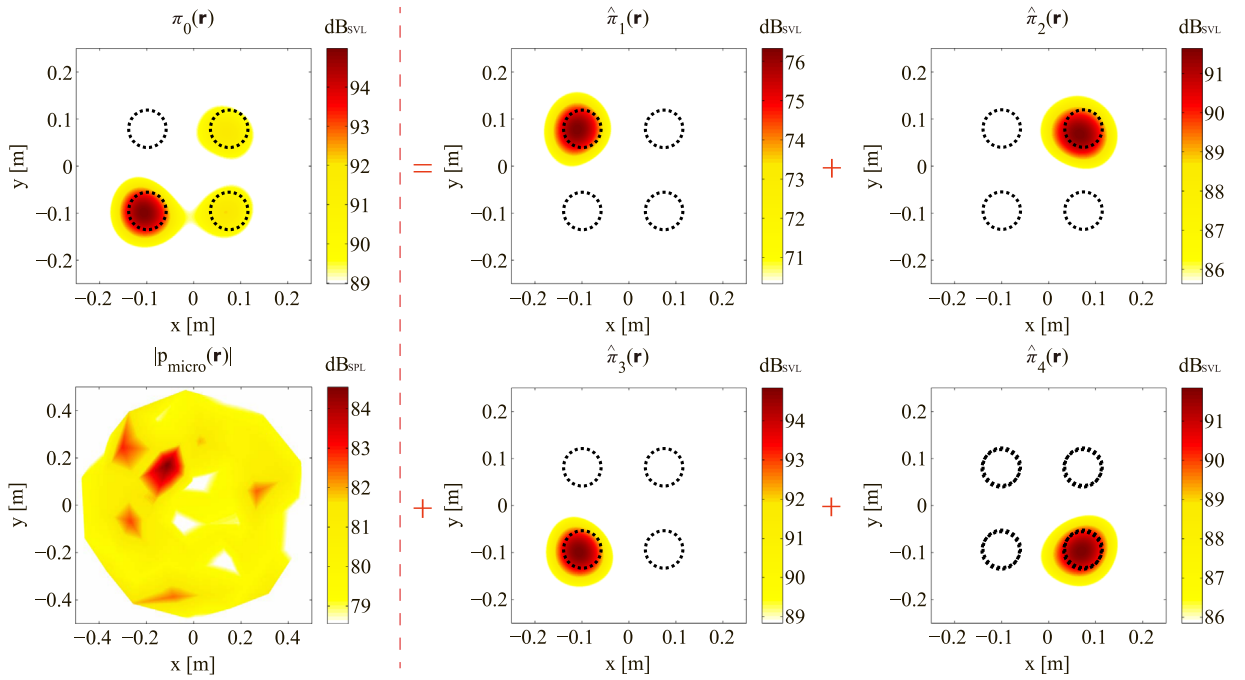

Fig. 20. Separated sources from joint statistical and spatial orthogonalization in the far-field at 2437 Hz. The total quadratic velocity $\pi_0(\mathbf{r},\omega)$ is decomposed into components $\hat{\pi}_i(\mathbf{r})$, $i = 1,\ldots,4$. (Dynamic range = 6 dB).

λ/L_s ratios. It can be seen in Fig. 26(a) and (b) that results obtained with PCA (i.e. statistical orthogonalization only) do not allow to separate the two uncorrelated source fields. Virtual sources do not coincide with the true sources. The first virtual source, see Fig. 26(a), embodies the two sources while the second virtual source has a spatial distribution similar to a dipole (in fact, by looking at the phase between the two lobes, it was seen that they are out-of-phase). Care must be taken thus, for the interpretation of results obtained with PCA for spatially distributed and closely spaced uncorrelated sources. On the

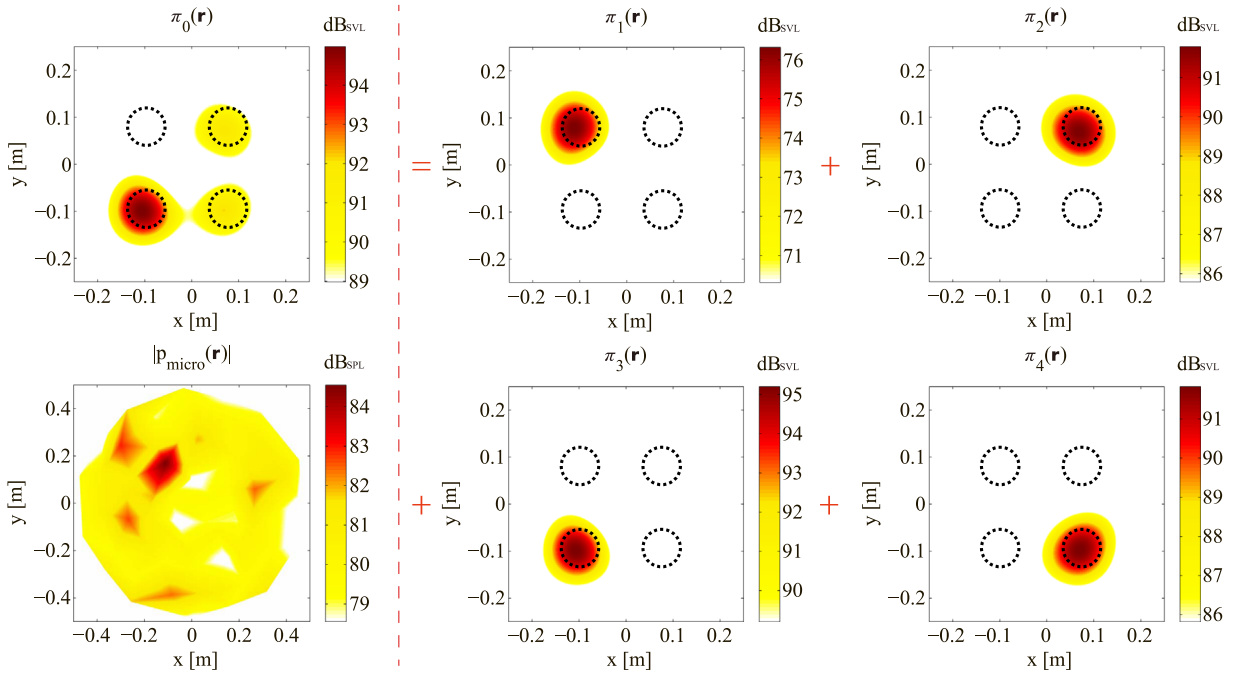

Fig. 21. True sources obtained from individual measurements in the far-field at 2437 Hz. (Dynamic range = 6 dB).

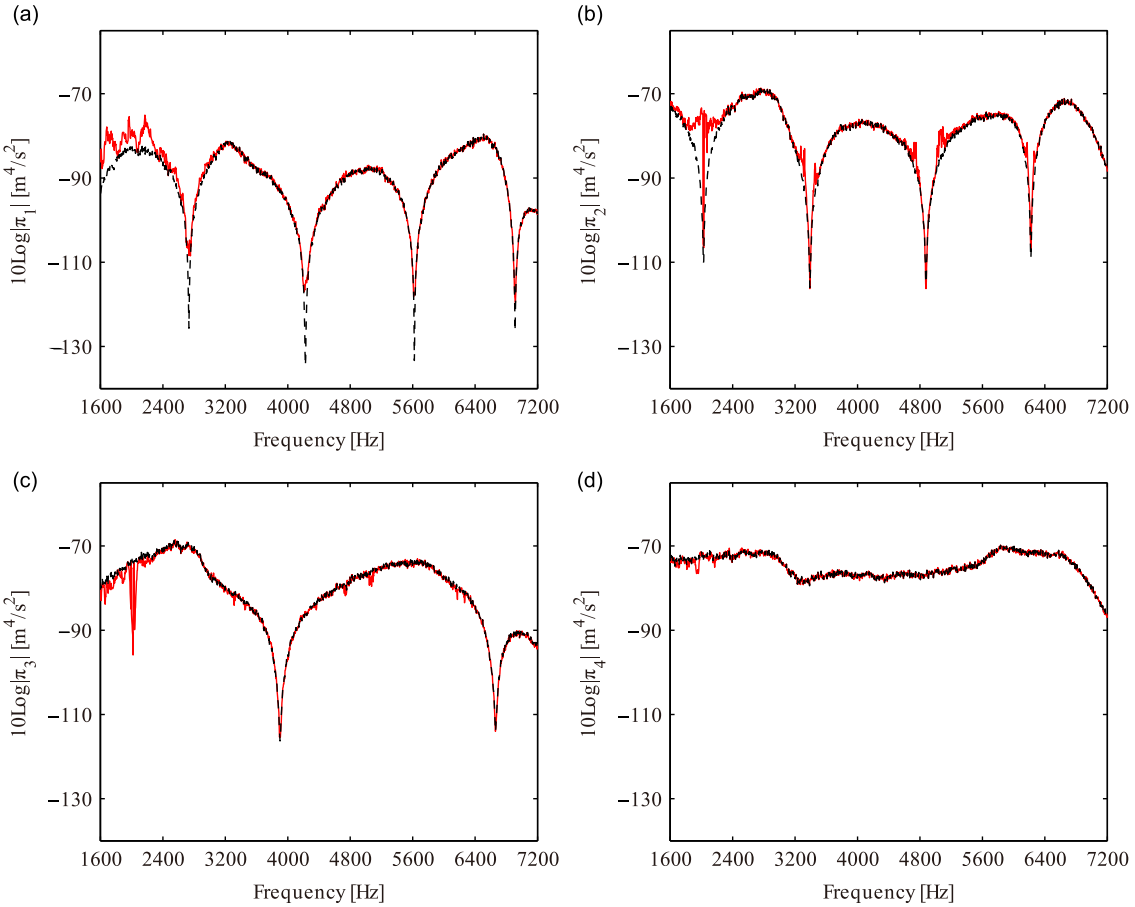

Fig. 22. Quadratic velocity spectra of the separated sources, (a) $\hat{\pi}_1(\omega)$, (b) $\hat{\pi}_2(\omega)$, (c) $\hat{\pi}_3(\omega)$, and (d) $\hat{\pi}_4(\omega)$, from joint statistical and spatial orthogonalization (red solid line) compared to true sources (black dashed line) in the far-field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

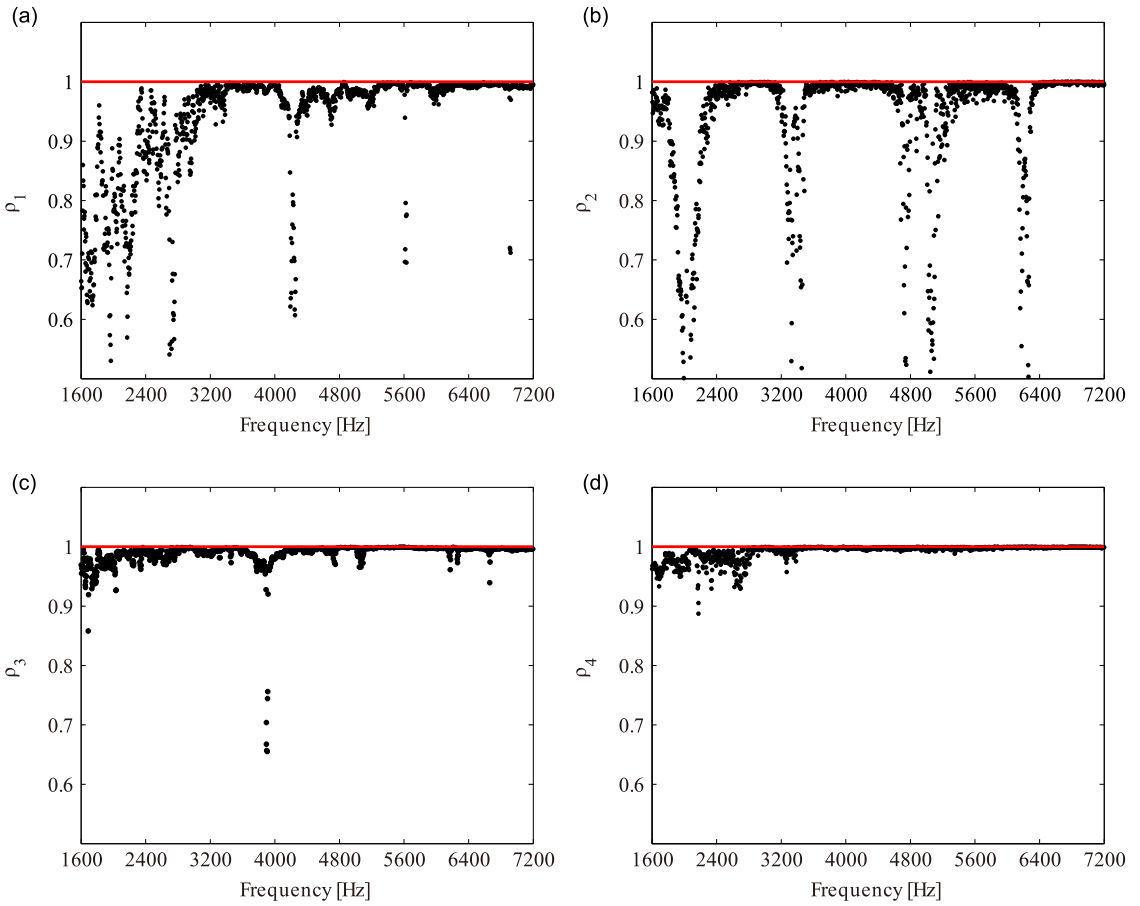

Fig. 23. Spatial correlation spectra between the separated and true sources, a) $\rho_1(\omega)$, b) $\rho_2(\omega)$, c) $\rho_3(\omega)$, and d) $\rho_4(\omega)$, in the far-field.

Table 2
Spatial correlation coefficients between the separated and true sources from SO and 2SO in the far-field at 2437 Hz.

	s_1	s_2	s_3	s_4
SO	1	0.86	1	0.84
2SO	0.97	1	0.99	0.98

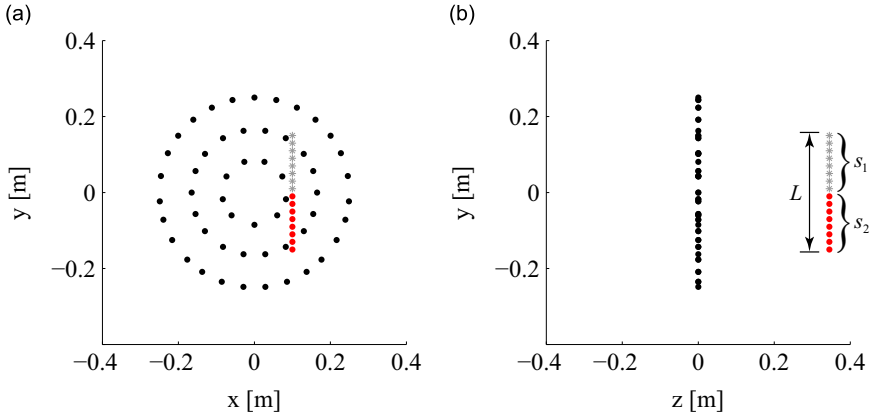

Fig. 24. Geometry of the simulation for the case of aeroacoustic sources: (a) front view and (b) side view. The microphone positions are shown as black dots. The source plane is located at a distance of 35 cm from the microphone plane. The sources are divided into two groups of mutually correlated sources (one group represented by gray stars and the other by red dots). Sources of different groups are uncorrelated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

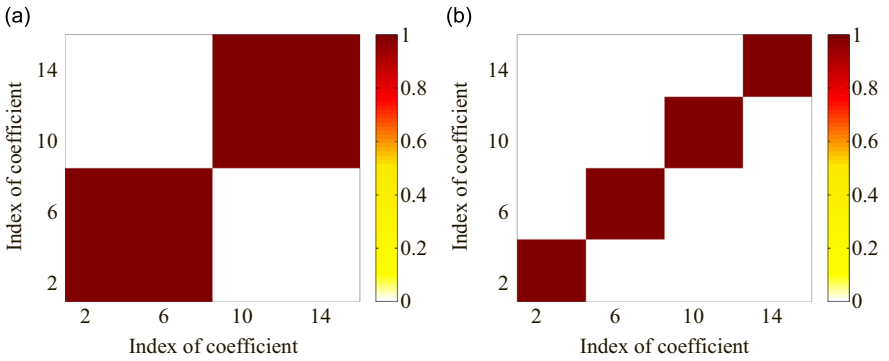

Fig. 25. Matrices of correlation coefficients correspond to the two simulation cases. 16 source coefficients are used for both cases: (a) two mutually uncorrelated groups and (b) four mutually uncorrelated groups. The value of correlation coefficient is given by the color scale and ranges from 0 (no correlation) to 1 (full correlation). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

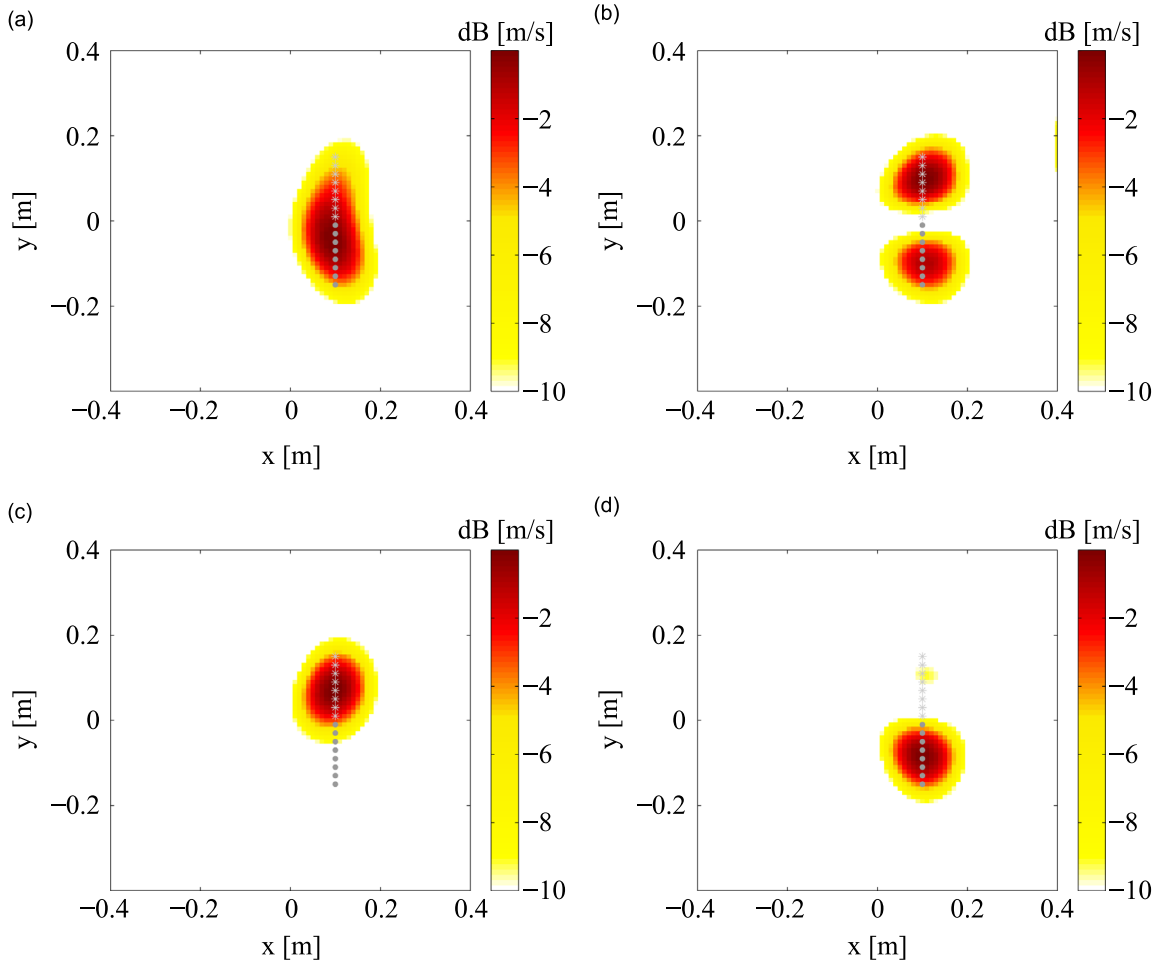

Fig. 26. Reconstructed source maps obtained with SO (the first row) and 2SO (the second row). Results are shown at 1500 Hz or $\lambda/L_s = 1.5$: (a) the first component or virtual source for SO; (b) the second component or virtual source for SO; (c) the first separated source for 2SO; and (d) the second separated source for 2SO.

other hand, results are improved if spatial orthogonality in addition to statistical orthogonality is forced, see panels (c) and (d) of Figs. 26–28. It can be seen that the uncorrelated source components are correctly separated at the studied frequencies.

Results for wavelengths of comparable size and smaller to the source dimensions, see Figs. 27 and 28, are basically the same as the previous. It can be seen that PCA results are slightly improved as the frequency increases, nevertheless it still does not provide a full separation of the two uncorrelated sources.

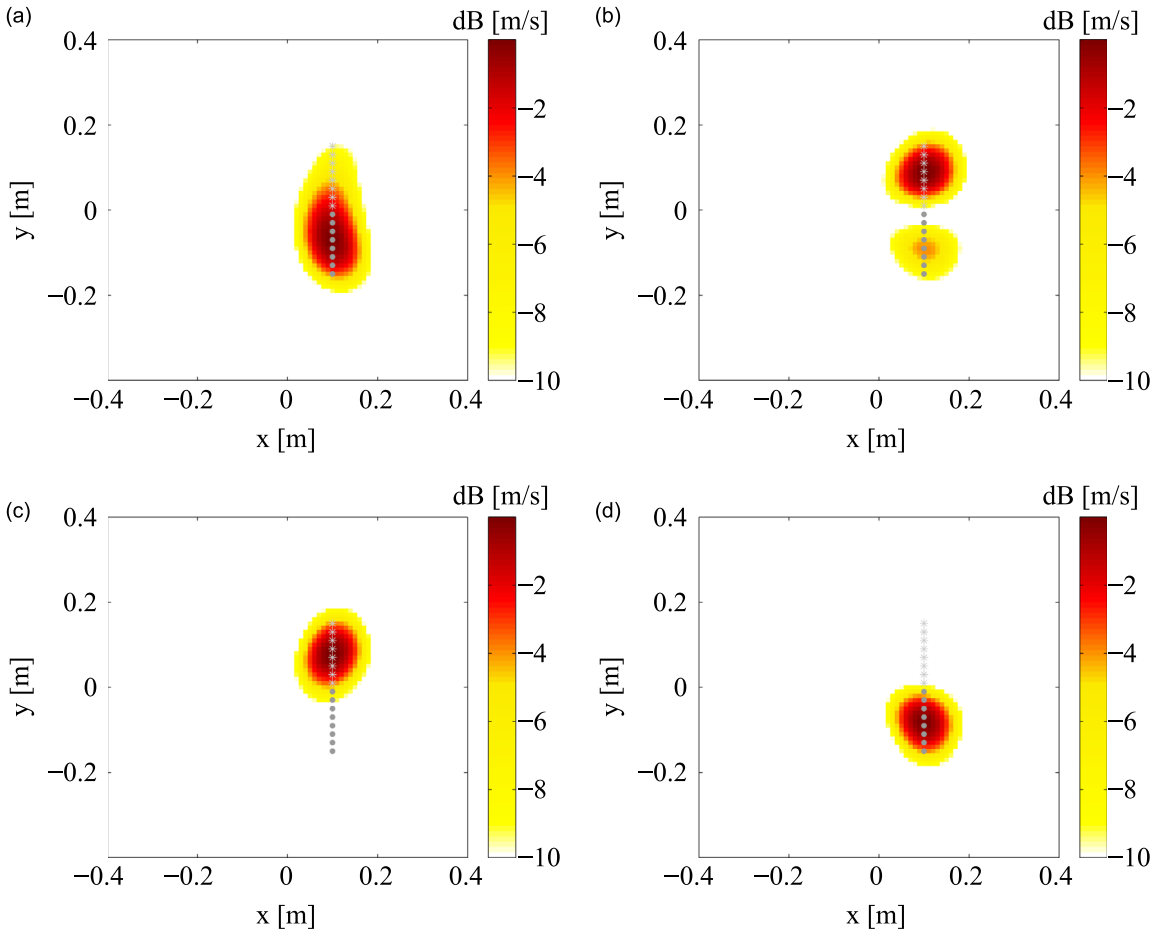

Fig. 27. Reconstructed sources obtained with SO (the first row) and 2SO (the second row). Results are shown at 2250 Hz or $\lambda/L_s = 1$: (a) the first component or virtual source; (b) the second component or virtual source; (c) the first separated source; and (d) the second separated source.

A second simulation experiment is realized with now four groups of uncorrelated sources as shown in Fig. 29.

The results obtained with PCA and with the proposed BSS algorithm are shown in Figs. 30–32. Similarly to the previous simulation case, we observe that statistical orthogonalization only (PCA) is not sufficient for the separation of the four uncorrelated components, even as frequency increases. The reconstructed source map corresponding to each virtual source does not correspond to the physical sources in this case. Contrary to the results obtained with SO, we can see that forcing the spatial orthogonality allows the correct separation of the four uncorrelated sources, see the second row of Figs. 30–32.

One may notice that at the frequency of 2250 Hz, the fourth separated source using 2SO, see Fig. 30(h), has still a small contribution from the other three sources. This is in fact due to the limited spatial resolution of the backpropagation (preliminary step to the BSS algorithm) method at this frequency. Note that the wavelength (λ) is large as compared to the source dimension $\lambda/L_s = 2$ at this frequency. This result could be possibly improved by using sparsity enforcing methods in the backpropagation step, however it is not in the scope of the present paper and it is left as a perspective of further research.

5.4.2. Real experiment in an anechoic wind-tunnel

The aeroacoustic experiment was carried out in an open-jet anechoic wind tunnel at the École Centrale de Lyon (ECL), see Fig. 33. A symmetric NACA 0012 airfoil with a chord length of 10 cm is placed at the potential core of a rectangular jet of width 15 cm and height 30 cm. A commercially available microphone array (HDcam) with 54 microphones (GRAS ¼" Type 40PL) is placed parallel to the flow direction at a distance of 35 cm from the NACA0012 airfoil, see Fig. 33(b). A turbulence grid placed inside the nozzle provides a turbulent flow condition impacting the airfoil, see Fig. 33(c). Measurements have been carried out for different free stream velocities and different flow regimes. A 56-channel SCADAS acquisition system is used with a sampling frequency of 25.6 kHz and the acquisition time set to 10 s.

It is well known that different noise mechanisms simultaneously appear when an airfoil is submitted to a turbulent flow [56–59]. One example is the noise originated from the impingement of upstream turbulence at the leading edge (known as turbulence-interaction noise or leading edge noise). Another one is the noise due to the scattering of the turbulent boundary

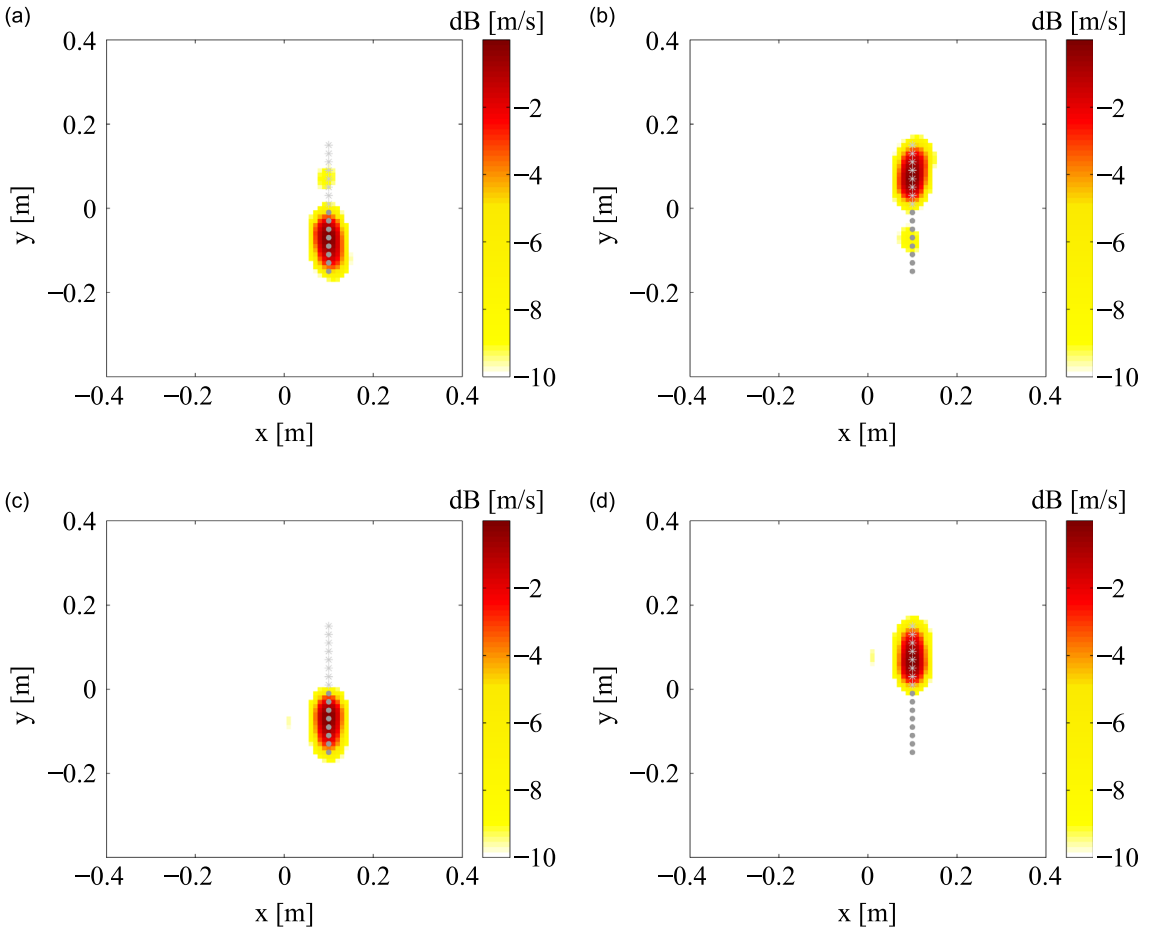

Fig. 28. Reconstructed sources obtained with SO (the first row) and 2SO (the second row). Results are shown at 4550 Hz or $\lambda/L_s = 0.5$: (a) the first component or virtual source; (b) the second component or virtual source; (c) the first separated source; and (d) the second separated source.

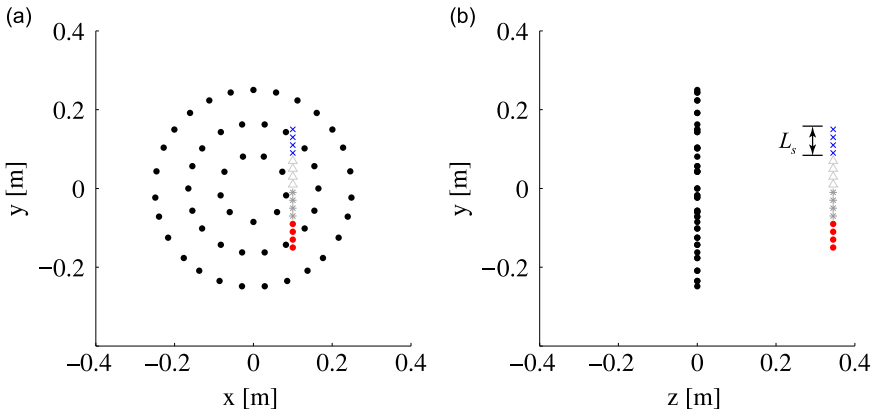

Fig. 29. Geometry of the simulation for the case of aeroacoustic sources: (a) front view and (b) side view. Microphone positions are shown as black dots. The source plane is located at 35 cm from the microphone plane. The source is divided into four groups of mutually correlated sources, each group with characteristic dimension L_s. Each two source groups are uncorrelated to each other.

layer as it passes through the trailing edge (trailing edge noise or airfoil self-noise). In practice for a turbulent flow, both of these noise mechanisms co-exist and the noise captured by a far field microphone will be a sum of both contributions. An example of measured acoustic pressure spectrum in Fig. 34 shows the broadband nature of the noise.

As an example of source reconstruction, Bayesian focusing is used for the backpropagation of measurements to the source surface. The reconstructed source field at different working frequencies is shown in Fig. 35. It can be seen that

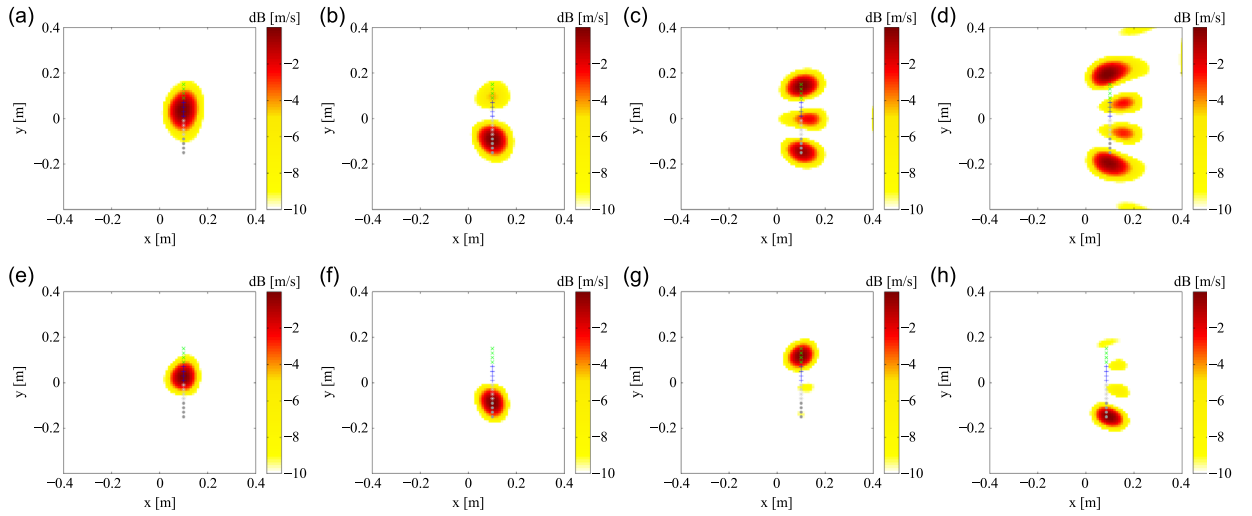

Fig. 30. Reconstructed sources obtained with SO (the first row) and 2SO (the second row). Results are shown at 2250 Hz or $\lambda/L_s = 2$: (a)–(d) the first to the fourth principal component or virtual source; and (e)–(h) the first to the fourth separated source.

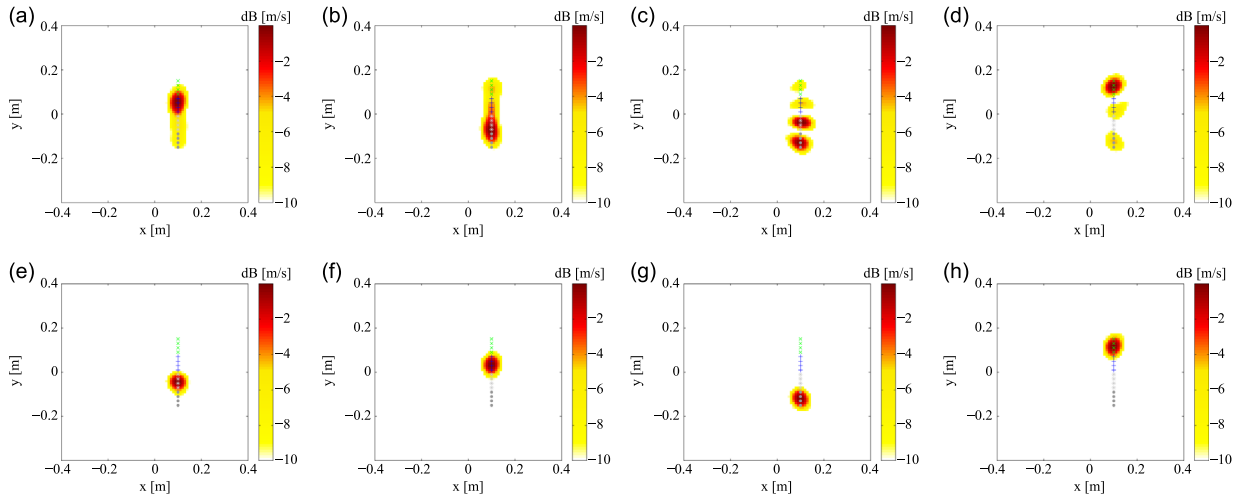

Fig. 31. Reconstructed sources obtained with SO (the first row) and 2SO (the second row). Results are shown at 4550 Hz or $\lambda/L_s = 1$: (a)–(d) the first to the fourth principal component or virtual source; and (e)–(h) the first to the fourth separated source.

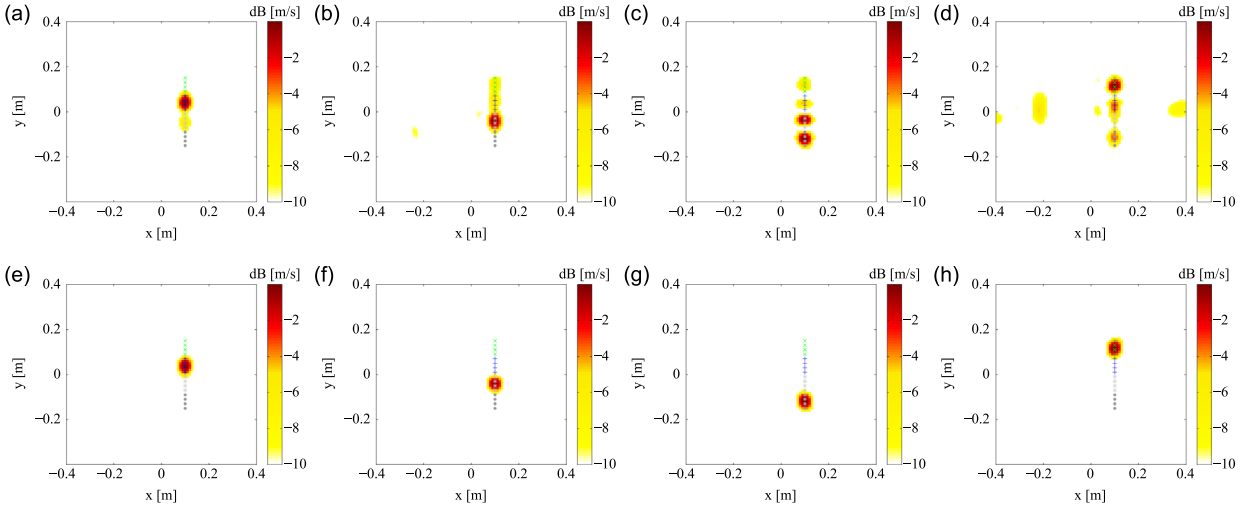

Fig. 32. Reconstructed sources obtained with SO (the first row) and 2SO (the second row). Results are shown at 5650 Hz or $\lambda/L_s = 0.8$: (a)–(d) the first to the fourth principal component or virtual source; and (e)–(h) the first to the fourth separated source.

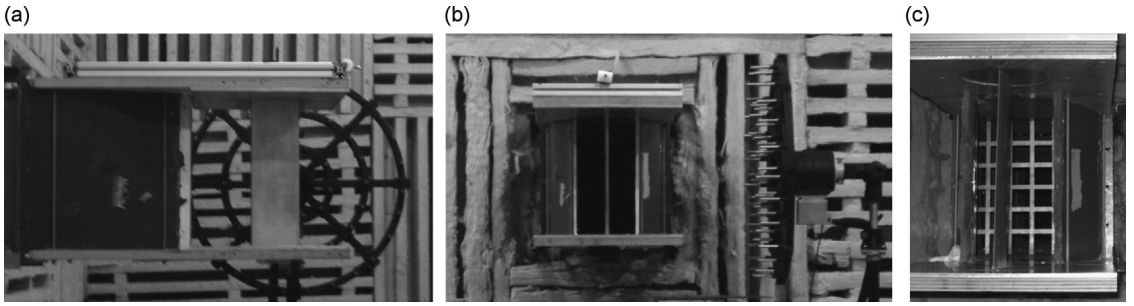

Fig. 33. Experimental set-up shows the open-jet anechoic wind tunnel: (a) side view, (b) front view, and (c) detail of turbulence grid installed upstream of the NACA 0012 airfoil.

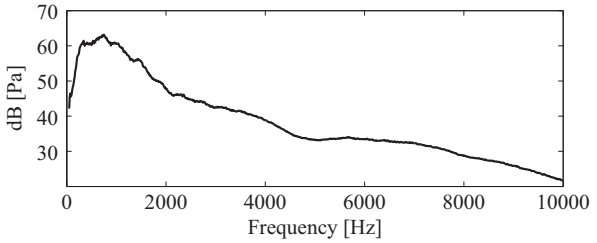

Fig. 34. Acoustic pressure spectrum averaged over all microphones for a free stream velocity of $U = 40$ m/s.

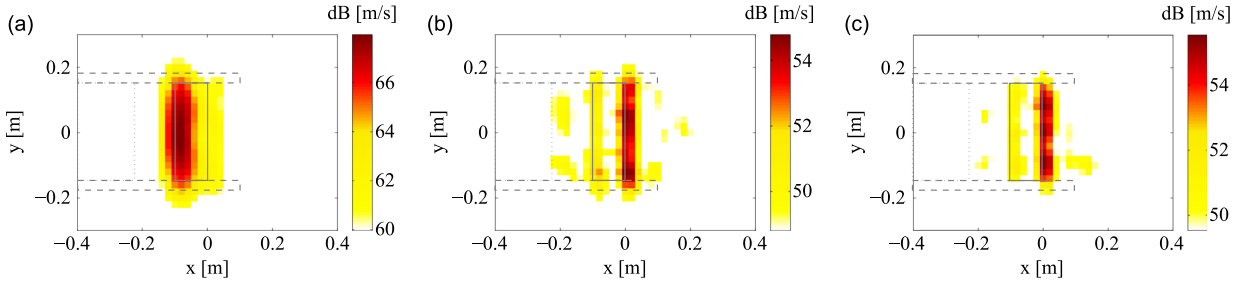

Fig. 35. Reconstructed normal velocity on the plane containing the airfoil at different frequencies: (a) 3100 Hz, (b) 5200 Hz, and (c) 5600 Hz. The airfoil geometry, the side plates and the nozzle exit are also sketched on all panels of the figure.

sources are basically located either on the leading edge or on the trailing edge of the airfoil, with different intensities depending on the frequency. At lower frequencies sources located at the leading edge are more energetic whereas at higher frequencies noise originated in the vicinity of trailing edge contributes the most to the far field acoustic pressure.

A first attempt has been to perform source separation by forcing the statistical orthogonality only, by means of PCA. Results of separated sources at the frequency of 3100 Hz are shown in Fig. 36. Similarly to the numerical experiments, SO fails to correctly separate the multiple uncorrelated sources. Indeed, this approach reconstructs "virtual sources" that, in this case, do not represent the physical behavior of true sources. Since the airfoil is submitted to a turbulent flow, one should rather expect physical sources to be spatially orthogonal along both the trailing edge and the leading edge.

Blind source separation is then applied by forcing statistical and spatial orthogonality (using the algorithm presented in Subsection 4.3). The separated sources at the two different frequencies are shown in Figs. 37 and 38. It can be noticed that uncorrelated sources are separated along the spanwise direction and are basically located either on the leading or the trailing edge of the airfoil. It is interesting to remark that contrary to the "global" result presented in Fig. 35(a), the individual contributions due to leading edge (Fig. 37(a)–(d)) and trailing edge noise (Fig. 37(e)–(g)) are separated. Indeed, this observation is in agreement with the discussion in Ref. [56] that argues that, based on measurements of wall pressure fluctuations along the airfoil chord at different turbulent ranges, turbulence-interaction noise and leading edge noise are uncorrelated. Notice that sources are also located at the edge of the nozzle, see Fig. 37(h) and (i), in spite of being more than 10 dB lower in level.

The results at a higher frequency (5600 Hz) are shown in Fig. 38. Similar observations are found, where sources on both leading and trailing edge noise are separated. Advantages of the BSS algorithm applied to this problem as compared to the global source reconstruction result are described hereafter. First, looking at the global result at the frequency of 3100 Hz, see

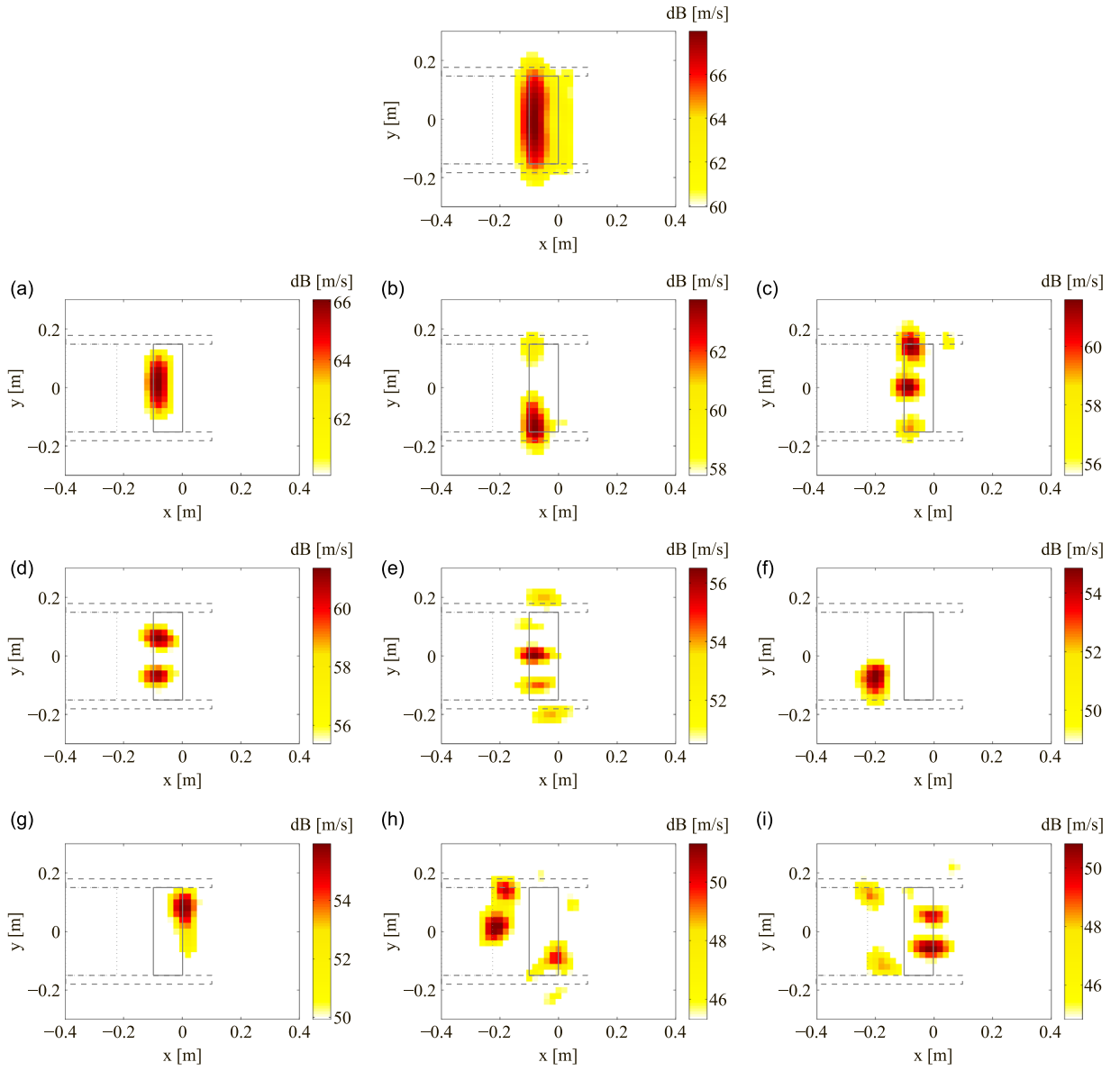

Fig. 36. Separated sources (a)–(i) obtained with statistical orthogonalization only. Results are shown at 3100 Hz. The dynamic range of results is set to 6 dB. The airfoil geometry, the side plates and the nozzle exit are also sketched on all panels of the figure.

Fig. 35(a), one sees that the separation between leading edge and trailing edge noises is limited by the dynamic range of the reconstructed source. One advantage of BSS is to improve the spatial separation capability of the reconstructed field, even if one source is considerably less energetic than others (see example in Subsection 5.3.2). Second, the uncorrelated sources separated by the method may give us an indication on the number of uncorrelated source components necessary to describe a physical phenomenon, equivalently to a source correlation length. For example, at the frequency of 5600 Hz, see Fig. 38, five uncorrelated components along the trailing edge are necessary to describe this noise source mechanism.

A final exercise is done to obtain the Power Spectral Density (PSD) related to each noise source mechanism, that is, trailing edge and leading edge noise. It consists of a post-processing of results delivered by the BSS algorithm. More precisely, each separated source is grouped by its spatial location and the source maps are spatially integrated around both the leading edge and the trailing edge. The obtained result is shown in Fig. 39. It can be seen that, as observed elsewhere [56], the turbulence-interaction noise is more important at low- and mid-frequencies whereas the trailing edge noise becomes dominant at higher frequencies. The same exercise has been tested directly with the results of the backpropagation step (i.e. without source separation), however the discrimination between leading edge and trailing edge noise was not successful, especially at low and mid-frequencies. In effect, the discrimination between both noise mechanisms was limited by the

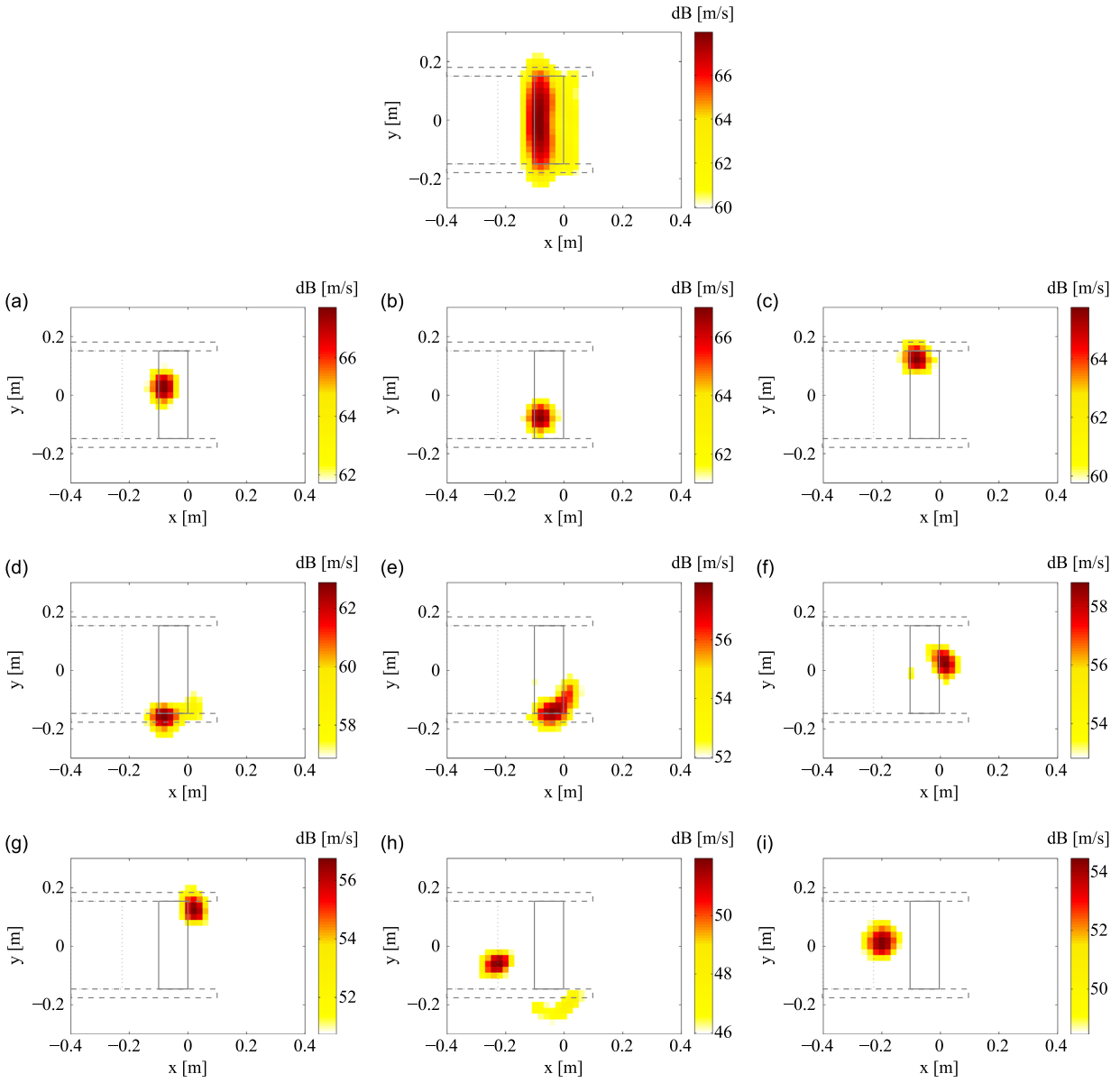

Fig. 37. Separated sources (a)–(i) obtained with the proposed BSS algorithm. Results are shown at 3100 Hz. The dynamic range of results is set to 6 dB. The airfoil geometry, the side plates and the nozzle exit are also sketched on all panels of the figure.

dynamic range of the reconstructed source field, in other words, residual contribution of leading edge noise was present in the integration region associated to trailing edge noise.

6. Conclusions

The findings of the paper are twofold. First, it addressed the question as whether PCA can return virtual sources that coincide with the true sound sources. It has been demonstrated that a unique solution exists in the case of "spatially orthogonal" incoherent sound sources. A particular case of interest of such a configuration is when sound sources are radiated from disjoint regions. Second, the endeavor for separating spatially disjoint sources has led to the so-called criterion of spatial orthogonalization and a related BSS algorithm. The proposed algorithm is fully original and has been found successful in situations where classical approaches, which do not make use of spatial information, fall short. It can accommodate any backpropagation method, such as beamforming, NAH, etc.

Two points are to be highlighted. First, the proposed BSS algorithm is computationally efficient and, because it runs about ten times faster, can advantageously replace the BSS method of Ref. [52] in the case of disjoint and compact sources;

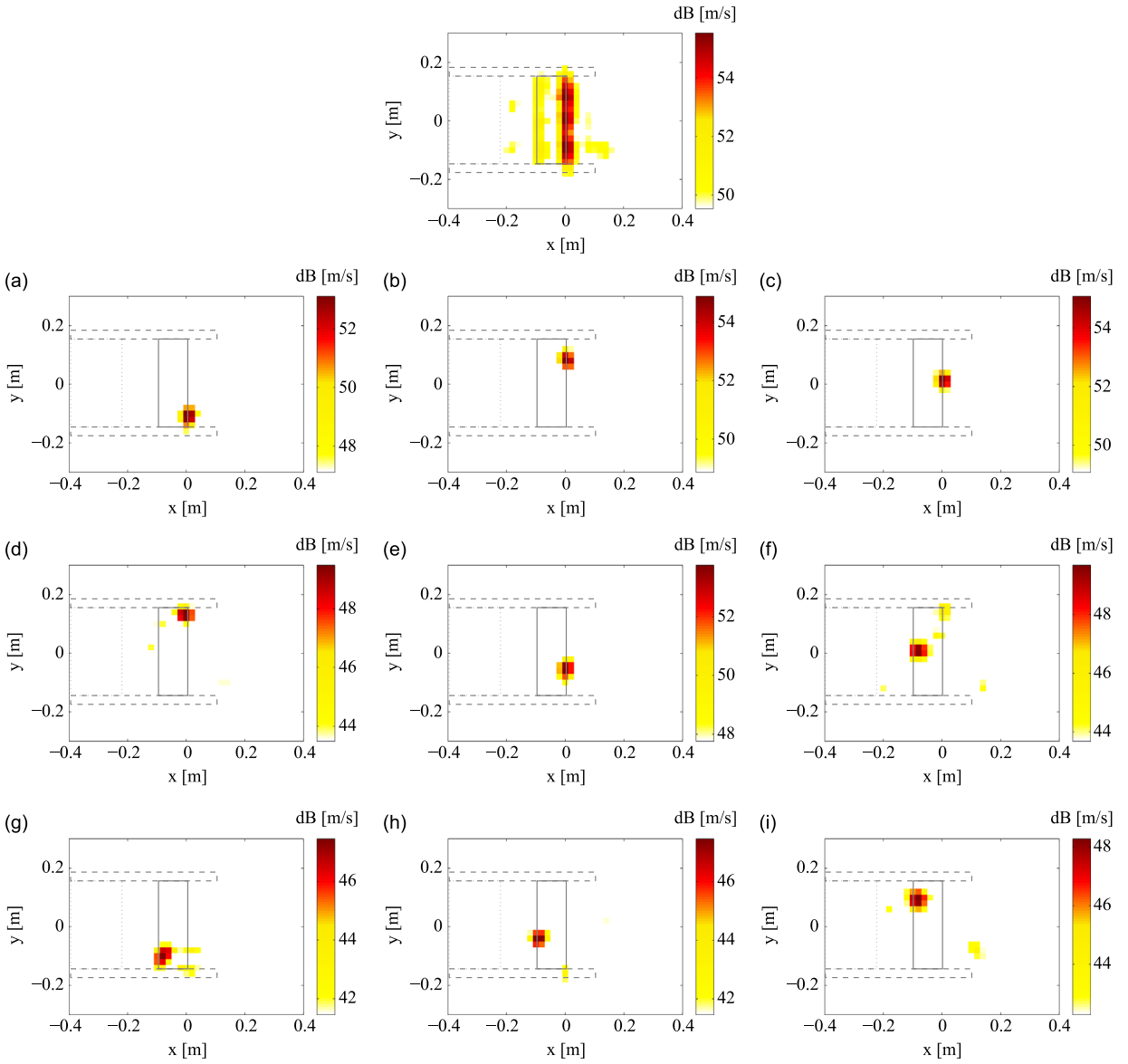

Fig. 38. Separated sources (a)–(i) obtained with the proposed BSS algorithm. Results are shown at 5600 Hz. The dynamic range of results is set to 6 dB. The airfoil geometry, the side plates and the nozzle exit are also sketched on all panels of the figure.

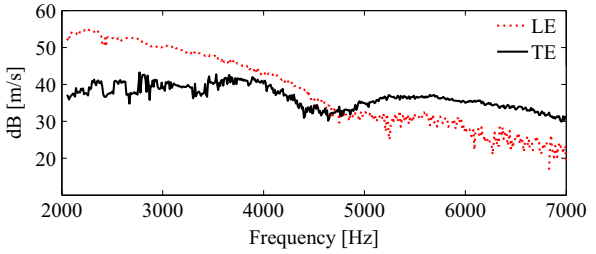

Fig. 39. Power spectral density of the particle velocity as obtained by the integration of the spatial source distribution around the Leading Edge (LE) and the Trailing Edge (TE).

however it should be carefully remembered that the principle of least spatial entropy in Ref. [52] applies more generally to sources which are "compact" and not necessarily disjoint. Second, the assumption of spatially disjoint sources strongly relies on the ability of backpropagation to resolve close sources, especially when spatial resolution is limited at low frequencies.

Finally, an interesting perspective for future work is to extend the proposed criteria of spatial orthogonalization to the case of coherent sources (i.e. mutually correlated), e.g. to be able to account for image sources produced by reflection. Even though this would imply skipping the PCA step, enough "equations" are theoretically left by spatial orthogonalization to solve the problem.

Acknowledgment

The authors are grateful to financial support from TUD COST Action TU 1105. They also express their thanks to the Institute of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg (FAU – Friedrich-Alexander-Universität Erlangen-Nürnberg), Germany. This work was conducted in the framework of the LabEx CeLyA (Centre Lyonnais d'Acoustique), (ANR-10-LABX-60/ANR-11-IDEX-0007) and the Labcom P3A (ANR-13-LAB2-0011-01), France. The experimental measurements of Subsection 5.4 have been carried out during the SEMAFOR project, supported by the FRAE.

References

[1] F. Fahy, J. Walker, *Advanced Applications in Acoustics, Noise and Vibration*, Spon Press, London, 2004.
[2] J.D. Maynard, E.G. Williams, Y. Lee, Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH, *Journal of Acoustical Society of America* 78 (1985) 1395–1413, http://dx.doi.org/10.1121/1.392911.
[3] E.G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, New York, 1999.
[4] J. Hald, Basic theory and properties of statistically optimized near-field acoustical holography, *Journal of Acoustical Society of America* 125 (2009) 2105–2120, http://dx.doi.org/10.1121/1.3079773.
[5] F. Jacobesen, V. Jaud, Statistically optimized near field acoustic holography using an array of pressure-velocity probes, *Journal of Acoustical Society of America* 121 (2007) 1550–1558, http://dx.doi.org/10.1121/1.2434245.
[6] S.F. Wu, On reconstruction of acoustic pressure fields using he Helmholtz equation least squares method, *Journal of Acoustical Society of America* 107 (2000) 2511–2522, http://dx.doi.org/10.1121/1.428639.
[7] A. Sarkissian, Extension of measurement surface in near-field acoustic holography, *Journal of Acoustical Society of America* 115 (2004) 1593–1596, http://dx.doi.org/10.1121/1.1645609.
[8] Q. Leclère, Acoustic imaging using under-determined inverse approaches: Frequency limitations and optimal regularization, *Journal of Sound and Vibration* 321 (2009) 605–619, http://dx.doi.org/10.1016/j.jsv.2008.10.022.
[9] J. Antoni, A. Bayesian, approach to sound source reconstruction: Optimal basis, regularization and focusing, *Journal of Acoustical Society of America* 131 (2012) 2873–2890, http://dx.doi.org/10.1121/1.3685484.
[10] M.B.S. Magalhães, R.A. Tenenbaum, Sound sources reconstruction techniques: A review of their evolution and new trends, *Journal of Acta Acustica united with Acustica* 90 (2004) 199–220.
[11] S.F. Wu, Methods for reconstructing acoustic quantities based on acoustic pressure measurements, *Journal of Acoustical Society of America* 124 (2008) 2680–2697, http://dx.doi.org/10.1121/1.2977731.
[12] J.S. Bendat, A.G. Piersol, *Random Data: Analysis and Measurement Procedures*, Wiley, New York, 1971.
[13] D. Hallman, J.S. Bolton, Multi-reference near-field acoustical holography Toronto, Canada, July, *Proceedings of Inter-noise* 92 (1992) 1165–1170.
[14] D.L. Hallman, J.S. Bolton, Comparison of multi-reference nearfield acoustical holography procedures, Proceedings of National Conference on Noise Control Engineering, Fort Lauderdale, USA, May 1994, pp. 929-934.
[15] J.F. Li, J.C. Pascal, C. Carles, Reconstruction of partially coherent sources by use of principal component analysis Newport Beach, USA, July, *Proceedings of Inter-noise* 95 (1995) 1355–1358.
[16] R.J. Ruhala, C.B. Burroughs, Separation of leading edge, trailing edge, and sidewall noise sources from rolling tires Ypsilanti, USA, April, *Proceedings of NOISE-CON* 98 (1998) 109–114.
[17] H.S. Kwon, Y.J. Kim, J.S. Bolton, Compensation for source nonstationarity in multireference, scan-based near-field acoustical holography, *Journal of the Acoustical Society of America* 113 (2003) 360–368, http://dx.doi.org/10.1121/1.1529669.
[18] B. Lafon, J. Antoni, M. Sidahmed, L. Polac, The concept of cyclic sound intensity and its application to acoustical imaging, *Journal of Sound and Vibration* 330 (2011) 2107–2121, http://dx.doi.org/10.1016/j.jsv.2010.11.003.
[19] M.A. Tomlinson, Partial source discrimination in near field acoustic holography, *Journal of Applied Acoustics* 57 (1999) 243–261, http://dx.doi.org/10.1016/S0003-682X(98)00058-9.
[20] K.U. Nam, Y.H. Kim, Visualization of multiple incoherent sources by the backward prediction of near-field acoustic holography, *Journal of the Acoustical Society of America* 109 (2001) 1808–1816, http://dx.doi.org/10.1121/1.1358888.
[21] Y.J. Kim, J.S. Bolton, H.S. Kwon, Partial sound field decomposition in multireference near-field acoustical holography by using optimally located virtual references, *Journal of Acoustical Society of America* 115 (2004) 1641–1652, http://dx.doi.org/10.1121/1.1642627.
[22] M.Y. Lee, J.S. Bolton, Scan-based near-field acoustical holography and partial field decomposition in the reference of noise and source level variation, *Journal of Acoustical Society of America* 119 (2006) 382–393, http://dx.doi.org/10.1121/1.2133717.
[23] K.U. Nam, Y.H. Kim, A partial field decomposition algorithm and its examples for near-field acoustic holography, *Journal of Acoustical Society of America* 116 (2004) 172–185, http://dx.doi.org/10.1121/1.1756896.
[24] S.M. Price, R.J. Bernhard, Virtual coherence: A digital signal processing technique for incoherent source identification, Proceedings of the Fourth International Modal Analysis Conference, Vol. 2, Los Angeles, USA, 1986, pp. 1256-1262.
[25] D. Otte, P. Sas, P. Vandeponseele, Principal component analysis for noise source identification, Proceedings of the International Modal Analysis Conference IMAC 6, Schenectary, USA, 1988.
[26] D. Otte, K. Fyfe, P. Sas, J. Leuridan, Use of principal component analysis for dominant noise source identification, Proceedings of the International Conference on Advances in the Control and Refinement of Vehicle Noise, Birmingham, United Kingdom, 1988.
[27] H. Mermoz, Imagerie, corrélation et modèles, *Annales Des Télécommunications* 31 (1976) 17–36, http://dx.doi.org/10.1007/BF02999092.
[28] C. Colangeli, P. Chiariotti, K. Janssens, Uncorrelated noise sources separation using inverse beamforming, in: J. D. Clerck (Ed.), Experimental Techniques, Rotating Machinery, and Acoustics, Vol. 8, Proceedings of the 33rd IMAC, A Conference and Exposition on Structural Dynamics, Springer, 2015, pp. 59-70. http://dx.doi.org/10.1007/978-3-319-15236-3_5.
[29] P. Ruiz, J.L. Lacoume, Extraction of independent sources from correlated inputs a solution based on cumulants, Proceedings of the Workshop on Higher-Order Spectral Analysis, Vail, June 1989, pp. 146-151. http://dx.doi.org/10.1109/HOSA.1989.735286.
[30] J.L. Lacoume, P. Ruiz, Separation of independent sources from correlated inputs, *IEEE Transactions on Signal Processing* 40 (1992) 3074–3078, http://dx.doi.org/10.1109/78.175753.

[31] P. Comon, Independent component analysis: a new concept? *International Journal of Signal Processing* 36 (1994) 287–314, http://dx.doi.org/10.1016/0165-1684(94)90029-9.
[32] P. Comon, C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Application*, Academic Press, Oxford, 2010.
[33] J.F. Cardoso, Blind signal separation: statistical principles, *International Journal of Proceeding of the IEEE* 86 (1998) 2009–2025.
[34] X.R. Cao, R.W. Liu, General approach to blind source separation, *International Journal of IEEE Transactions on Signal Processing* 44 (1996) 562–571.
[35] T. Pham, Blind separation of instantaneous mixture of sources via the Gaussian mutual information criterion, *International Journal of Signal Processing* 81 (2001) 855–870.
[36] J. Hérault, C. Jutten, Space or time adaptive signal processing by neural networks models. International Conference on Neural Networks for Computing, Snowbird, 1986, Utah, USA, pp. 206-211.
[37] S. Bourennane, A. Bendjama, Locating wide band acoustic sources using higher order statistics, *Applied Acoustics* 63 (2002) 235–251, http://dx.doi.org/10.1016/S0003-682X(01)00039-1.
[38] J. Kociński, S. Drgas, E. Ozimek, Evaluation of blind source separation for different algorithms based on second order statistics and different spatial configurations of directional microphones, *Applied Acoustics* 73 (2012) 109–116, http://dx.doi.org/10.1016/j.apacoust.2011.06.016.
[39] O. Bunting, D. Chesmore, Time frequency source separation and direction of arrival estimation in a 3D soundscape environment, *Applied Acoustics* 74 (2013) 264–268, http://dx.doi.org/10.1016/j.apacoust.2011.05.018.
[40] L. Zhang, Y. Jiang, L. He, Bidirectional coupled noise sources separation over the same frequency band using convolutive blind separation method, *Applied Acoustics* 74 (2013) 301–306, http://dx.doi.org/10.1016/j.apacoust.2012.08.006.
[41] R. Brillinger, *Time Series: Data Analysis and Theory*, SIAM, Philadelphia, 2001.
[42] H. Buchner, R. Aichner and W. Kellermann, TRINICON: a versatile framework for multichannel blind signal processing, Proceedings of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, Montreal, Canada, May 2004, pp. 889-892. http://dx.doi.org/10.1109/ICASSP.2004.1326688.
[43] H. Saruwatari, S. Kurita, K. Takeda, Blind source separation combining frequency-domain ICA and beamforming, Proceedings of 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 5, Salt Lake City, USA, May 2001, pp. 2733-2736. http://dx.doi.org/10.1109/ICASSP.2001.940211.
[44] L.C. Parra, C.V. Alvino, Geometric source separation: merging convolutive source separation with geometric beamforming, *IEEE Transactions on Speech and Audio Processing* 10 (2002) 352–362, http://dx.doi.org/10.1109/TSA.2002.803443.
[45] N. Mitianoudis, M. E. Davies, Using beamforming in the audio source separation problem, in: Proceedings of the Seventh International Symposium on Signal Processing and Its Applications, Vol. 2, Paris, France, July 2003, pp. 89-92. http://dx.doi.org/10.1109/ISSPA.2003.1224822.
[46] S. Araki, S. Makino, Y. Hinamoto, R. Mukai, T. Nishikawa, H. Saruwatari, Equivalence between frequency-domain blind source separation and frequency-domain adaptive beamforming for convolutive mixtures, *EURASIP Journal on Applied Signal Processing* 11 (2003) 1157–1166, http://dx.doi.org/10.1155/S1110865703305074.
[47] T. Nishikawa, H. Saruwatari, K. Shikano, Blind separation of more than two sources based on high-convergence algorithm combining ICA and beamforming, in: Proceedings of the 13th European Signal Processing Conference, Antalya, Turkey, September 2005, pp. 1-4.
[48] K. Reindl, P. Prokein, E. Fischer, Y. Zheng, W. Kellermann, Combining monaural beamforming and blind source separation for binaural speech enhancement in multi-microphone hearing aids, in: Proceedings of ITG Conference on Speech Communication (Sprachkommunikation 2010), Bochum, Germany, October 2010.
[49] N. Ogasawara, T. Nishino, K. Takeda, Blind source separation based on acoustic pressure distribution and normalized relative phase using dodecahedral microphone array, in: Proceedings of the 17th European Signal Processing Conference, Glasgow, Scotland, August 2009, pp. 1413-1417.
[50] J. Thiemann, E. Vincent, An experimental comparison of source separation and beamforming techniques for microphone array signal enhancement, in: Proceedings of 2013 IEEE International Workshop on Machine Learning for Signal Processingsignal enhancement, Southampton, UK, September 2013, pp. 1-5. http://dx.doi.org/10.1109/MLSP.2013.6661961.
[51] E. Zhang, J. Antoni, B. Dong, H. Snoussi, Bayesian space-frequency separation of wide-band sound sources by a hierarchical approach, *Journal of Acoustical Society of America* 132 (2012) 3240–3250, http://dx.doi.org/10.1121/1.4754530.
[52] B. Dong, J. Antoni, E. Zhang, Blind separation of sound sources from the principle of least spatial entropy, *Journal of Sound and Vibration* 333 (2014) 2643–2668, http://dx.doi.org/10.1016/j.jsv.2013.12.011.
[53] L.L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley Publishing Company, Amsterdam, 1991.
[54] A. Cichocki, S. Amari, *Adaptive blind signal and image processing: Learning algorithms and applications*, Wiley, Chichester, 2002, 586.
[55] J.F. Cardoso, ⟨http://sig.enst.fr/~cardoso/stuff.html⟩ (accessed 13:26 14 October 2013).
[56] S. Moreau, M. Roger, Competing broadband noise mechanisms in low-speed axial fans, *AIAA Journal* 45 (2007) 48–57, http://dx.doi.org/10.2514/1.14583.
[57] M. Roger, S. Moreau, Extensions and limitations of analytical airfoil broadband noise models, *International Journal of Aeroacoustics* 9 (2010) 273–305.
[58] R.K. Amiet, Acoustic radiation from an airfoil in a turbulent stream, *Journal of Sound and Vibration* 41 (1975) 407–420, http://dx.doi.org/10.1016/S0022-460X(75)80105-2.
[59] T.F. Brooks, T.H. Hodgson, Trailing edge noise prediction from measured surface pressures, *Journal of Sound and Vibration* 78 (1981) 69–117, http://dx.doi.org/10.1016/S0022-460X(81)80158-7.