Blind Separation of Uncorrelated Sound Sources from the Principle of Least Spatial Complexity

Bin Dong, Jérôme Antoni

BLIND SEPARATION OF UNCORRELATED SOUND SOURCES FROM THE PRINCIPLE OF LEAST SPATIAL COMPLEXITY

DONG Bin and Jérôme ANTONI

Vibrations and Acoustic Laboratory, INSA of Lyon (National Institute for Applied Sciences), Building of St. Exupéry, 25 bis avenue of Jean Capelle, 69621 Villeurbanne Cedex, France.
e-mail: bin.dong@insa-lyon.fr

A novel algorithm based on the principle of least spatial complexity is proposed to decompose a sound field into uncorrelated sources by minimizing its spatial entropy. The global sound source distribution is first reconstructed from measurements returned by an array of microphones. Then it is transformed into uncorrelated virtual sources by means of the eigenvalue decomposition. However, it is well-known that having uncorrelated sources is a necessary but by no means a sufficient condition for separation. In order to find a unique solution with physical meaning, it is proposed to invoke the principle of least spatial complexity which forces an optimal solution with maximum spatial compactness. This is accomplished thanks to a novel conjugate gradient algorithm operating on the Stiefel manifold. An experiment is conducted to validate the proposed algorithm and evaluate its performance. Final results show that the sound field of each separated source perfectly matches that of the actual one, as would be obtained if all other sources could be switched off one by one.

1. Introduction

Localization and quantification of sound sources have attracted a lot of attention since many decades[1]. An ideal analysis procedure for identifying sound sources is as follows: sound sources stemming from distinct physical origins are decomposed into corresponding contributions based on measurements returned by an array of microphones. This is in essence a characteristic problem of blind source separation (BSS), which turns out extremely difficult to solve in practice. There have been numerous approaches to BSS in signal processing[2-6]. Unfortunately, most of these approaches can not be directly applied to sound source separation. One reason is the issue involves backpropagating sound waves from the microphone array to the source surface, which is a severely ill-posed inverse problem prone to produce extremely unstable results[7-8]. Another reason is that many BSS approaches require sources to be non-Gaussian (more exactly, no more than one Gaussian source) in order to properly define independence (independent sources are sought by making their joint probability distribution separable). This condition is hardly met in acoustics where sound propagation is often described in frequency domain. Indeed, Fourier Transformed signals are forced to tend to a complex Gaussian distribution by virtue of the central limit theorem (CLT), which precludes a unique definition of independence (a separable joint Gaussian probability distribution implies uncorrelated, but not independent sources, that is an infinite set of solutions versus a unique one to separation). For instance, that means independent component analysis (ICA), a popular non-

Gaussian based BSS methods, could not be applied for that reason. In order to overcome the limit of non-Gaussian, a series of BSS approaches employing the concept "information entropy" are proposed[9-13]. However, the sound sources are the signals of both frequency and space after doing FT. One peculiarity which singles out the acoustical context, however, is that the effort to reconstruct a spatial distribution suggests exploiting spatial information rather than probabilistic properties (i.e. construing independence through spatial properties instead of a probabilistic definition). This led the authors to propose the concept of "spatial entropy", inspired by "Shannon entropy", which quantifies the spatial complexity of a source distribution. Uncorrelated sources are then separated that make the final distribution as simple as possible, that is which minimize the spatial entropy.

The rest of the paper is organized as follows: the global reconstruction of sound source distribution is introduced first. This distribution is then decomposed into uncorrelated "virtual" sources which, however, do not enjoy the property of uniqueness. The principle of least spatial complexity is then employed to find a unique solution from the set of virtual sources. Finally, the validity of the proposed algorithm is demonstrated by a laboratory experiment. Properties of the proposed algorithm are then discussed in terms of final results.

2. Reconstructing the sound source distribution

The aim of this section is to introduce the direct propagation problem together with the corresponding to be used hereafter. Let us consider the sound field produced by a combination of n_s uncorrelated sound sources locates in free space and the resulting pressure signals recorded at some distance by an array of M microphones (see Fig. 1).

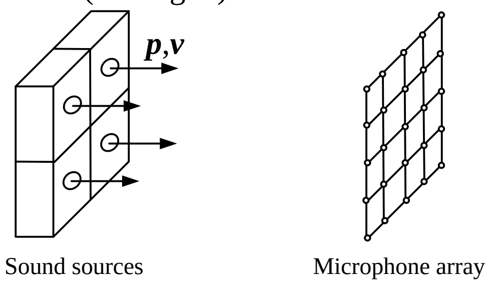

Figure 1. Illustration of a sound field produced from uncorrelated sources recorded by a microphone array

The M acoustic signals recorded by the array are denoted by vector $\mathbf{p}$, whose each element reads $p(\mathbf{r}_\mathrm{m},t)$ at time t, with $\mathbf{r}_\mathrm{m}$ the position vector of the m-th microphone. The pressure signal $p(\mathbf{r}_\mathrm{m},t)$ is further divided into a series of snapshots $p(\mathbf{r}_m, t; \varpi)$, where ϖ stands for the snapshot label. Every snapshot $p(\mathbf{r}_m, t; \varpi)$ is then Fourier transformed from the time domain to the frequency domain and denoted by $p(\mathbf{r}_m, \omega; \varpi)$, with ω the radial frequency.

The central assumption of this work is that the pressure field is produced by n_s uncorrelated sound sources (e.g. normal component of particle velocity, parietal pressure) $s_i(\mathbf{r}, \omega; \varpi)$, i=1, 2, …, n_s, such that

$$p(\mathbf{r}_m, \omega; \varpi) = \int_\Gamma \mathbf{G}(\mathbf{r}_m \,\|\, \mathbf{r}, \omega) \sum_{i=1}^{n_s} s_i(\mathbf{r}, \omega; \varpi) d\Gamma(\mathbf{r}), \tag{1}$$

where $\mathbf{r}$ is the position vector of the sound source, $\mathbf{G}(\mathbf{r}_m \,\|\, \mathbf{r}, \omega)$ is the Green function in free space and Γ is the sound source surface. The next step is now to discretize the source surface Γ. Let replace the position vector $\mathbf{r}$ by N samples $\mathbf{r}_l$, l = 1, 2, …, N, associated with a related small surface element $\Delta\Gamma_l$. The discrete form of Eq. (1) is then

$$p(\mathbf{r}_m, \omega; \varpi) = \sum_{l=1}^{N} \mathbf{G}(\mathbf{r}_m \,\|\, \mathbf{r}_l, \omega) \sum_{i=1}^{n_s} s_i(\mathbf{r}_l, \omega; \varpi) \Delta\Gamma_l, \tag{2}$$

which could be rewritten in a matrix form as:

$$\mathbf{p}(\varpi)=\begin{bmatrix} p(\mathbf{r}_1,\omega;\varpi) \\ \vdots \\ p(\mathbf{r}_M,\omega;\varpi) \end{bmatrix} = \begin{bmatrix} G(\mathbf{r}_1 \| \mathbf{r}_1,\omega)\Delta\Gamma_1 & \cdots & G(\mathbf{r}_1 \| \mathbf{r}_N,\omega)\Delta\Gamma_N \\ \vdots & \ddots & \vdots \\ G(\mathbf{r}_M \| \mathbf{r}_1,\omega)\Delta\Gamma_1 & \cdots & G(\mathbf{r}_M \| \mathbf{r}_N,\omega)\Delta\Gamma_N \end{bmatrix} \begin{bmatrix} s(\mathbf{r}_1,\omega;\varpi) \\ \vdots \\ s(\mathbf{r}_N,\omega;\varpi) \end{bmatrix} = \mathbf{G}\mathbf{s}(\varpi)\,. \tag{3}$$

In Eq. (3), $s(\mathbf{r}_l,\omega;\varpi)$ is the sum of the n_s uncorrelated sound sources $s_i(\mathbf{r}_l,\omega;\varpi)$ under the *l*-th position vector $\mathbf{r}_l$; as suggested in Ref. [14], it is conveniently represented by an optimal basis $\mathbf{\Phi}$ with a related vector of coefficients $\mathbf{c}$:

$$\mathbf{s}(\varpi)=\begin{bmatrix} s(\mathbf{r}_1,\omega;\varpi) \\ \vdots \\ s(\mathbf{r}_N,\omega;\varpi) \end{bmatrix} = \begin{bmatrix} \Phi_1(\mathbf{r}_1,\omega) & \cdots & \Phi_K(\mathbf{r}_1,\omega) \\ \vdots & \ddots & \vdots \\ \Phi_1(\mathbf{r}_N,\omega) & \cdots & \Phi_K(\mathbf{r}_N,\omega) \end{bmatrix} \begin{bmatrix} c_1(\omega;\varpi) \\ \vdots \\ c_K(\omega;\varpi) \end{bmatrix} = \mathbf{\Phi}\mathbf{c}(\varpi)\,. \tag{4}$$

where the optimal basis dimension, *K*, is equal to the number of microphones in the array. The recovery of the unknown vector of coefficient, $\mathbf{c}(\varpi)$, from the measured pressure signals $\mathbf{p}(\varpi)$, is then obtained as

$$\mathbf{c}(\varpi)=\left[\mathbf{G}\mathbf{\Phi}\right]^{+}\mathbf{p}(\varpi)\,, \tag{5}$$

where the exact expression of the pseudo-inverse $\left[\mathbf{G}\mathbf{\Phi}\right]^{+}$ is detailed in Ref. [14] and not reported here due to lack of space. This surely solves the source *reconstruction* problem, but not source *separation*; in other words, coefficients $c_{ki}(\omega;\varpi)$, k = 1, 2, …, *K*, assigned to the *i*-th sound source $s_i(\mathbf{r},\omega;\varpi)$ still remain unknown. The object of the paper is specifically to find every uncorrelated sound source $s_i(\mathbf{r},\omega;\varpi)$ and its corresponding coefficient vector $\mathbf{c}_i(\omega;\varpi)$, or in a matrix form, to solve:

$$\mathbf{S}(\varpi)=\begin{bmatrix} s_1(\mathbf{r}_1,\omega;\varpi) & \cdots & s_{n_s}(\mathbf{r}_1,\omega;\varpi) \\ \vdots & \ddots & \vdots \\ s_1(\mathbf{r}_N,\omega;\varpi) & \cdots & s_{n_s}(\mathbf{r}_N,\omega;\varpi) \end{bmatrix} = \begin{bmatrix} \Phi_1(\mathbf{r}_1,\omega) & \cdots & \Phi_K(\mathbf{r}_1,\omega) \\ \vdots & \ddots & \vdots \\ \Phi_1(\mathbf{r}_N,\omega) & \cdots & \Phi_K(\mathbf{r}_N,\omega) \end{bmatrix} \begin{bmatrix} c_{11}(\omega;\varpi) & \cdots & c_{1n_s}(\omega;\varpi) \\ \vdots & \ddots & \vdots \\ c_{K1}(\omega;\varpi) & \cdots & c_{Kn_s}(\omega;\varpi) \end{bmatrix} = \mathbf{\Phi}\mathbf{C}(\varpi)\,. \tag{6}$$

In order to do so, let us first note that the n_s coefficients $c_{ki}(\omega;\varpi)$ attached to the same basis function $\mathbf{\Phi}_k$ are necessarily uncorrelated, since the sound sources $s_i(\mathbf{r},\omega;\varpi)$ are. Therefore, combing Eqs. (4) and (6), $c_k(\omega;\varpi)$ can be further expanded as

$$c_k(\omega;\varpi)=\sum_{i=1}^{n_s} c_{ki}(\omega;\varpi)=\sum_{i=1}^{n_s} a_{ki}(\omega)\mathbb{S}_i(\omega;\varpi)\,, \tag{7}$$

where $\mathbb{S}_i(\omega;\varpi)$ is a "latent variable" which completely describes the probabilistic properties of the *i*-th source, but does not depend on the index *k* of spatial basis $\mathbf{\Phi}$, and $a_{ki}(\omega)$ corresponds to the coefficient of the *i*-th latent variable in the decomposition of the *k*-th coefficient $c_{ki}(\omega;\varpi)$. Without loss of generality, all latent variables will be assigned unit power. Equation (5) can now be rewritten more compactly as:

$$\mathbf{c}(\varpi) = \begin{bmatrix} c_1(\omega;\varpi) \\ \vdots \\ c_K(\omega;\varpi) \end{bmatrix} = \begin{bmatrix} a_{11}(\omega)\cdots & a_{1n_s}(\omega) \\ \vdots \quad \ddots & \vdots \\ a_{K1}(\omega)\cdots & a_{Kn_s}(\omega) \end{bmatrix} \begin{bmatrix} \mathbb{S}_1(\omega;\varpi) \\ \vdots \\ \mathbb{S}_{n_s}(\omega;\varpi) \end{bmatrix} = \mathbf{A}\mathbb{S}(\varpi) \,. \tag{8}$$

After ignoring the dependence on ϖ for notational simplicity, one has

$$\mathbf{p} = \mathbf{Gs} = \mathbf{G\Phi A}\mathbb{S} \,, \tag{9}$$

where the parameters $\mathbf{p}$, $\mathbf{G}$ and $\mathbf{\Phi}$ are all known quantities and parameters $\mathbf{A}$ and $\mathbb{S}$ are the unknowns to be determined. The next subsection explains how to solve this problem.

3. Pre-whitening of measurements

In order to find the unknown coefficient matrix **A**, its singular value decomposition (SVD) is considered first

$$[\mathbf{A}]_{SVD} = \mathbf{UDV}^H \,, \tag{10}$$

where **U** and **V** are unitary matrices, **D** is a nonnegative real diagonal matrix and superscript 'H' stands for the Hermitian transpose.

Since the n_s sound sources are mutually uncorrelated, the latent variables $\mathbb{S}_i(\omega;\varpi)$ (which are assumed of unit power without loss of generality) are seen to have a diagonal correlation matrix

$$\mathbb{E}\left\{\mathbb{S}\mathbb{S}^H\right\} = \mathbb{I} \,, \tag{11}$$

with $\mathbb{E}$ the expectation operator (ensemble average over all snapshots ϖ).

Therefore, according to Eqs. (9) - (11), one has

$$\mathbb{E}\left\{\mathbf{pp}^H\right\} = \mathbb{E}\left\{\mathbf{G\Phi A}\mathbb{S}\mathbb{S}^H\mathbf{A}^H\mathbf{\Phi}^H\mathbf{G}^H\right\} = \mathbf{G\Phi UD}^2\mathbf{U}^H\mathbf{\Phi}^H\mathbf{G}^H \,, \tag{12}$$

which uniquely returns the matrices **U** and **D** from the eigenvalue decomposition (EVD) of the correlation matrix $\mathbb{E}\left\{\mathbf{pp}^H\right\}$, provided that $\mathbf{G\Phi}$ is known. However, the unitary matrix **V** which enters the SVD of matrix **A** is still missing. This proves that there exists an infinite number of uncorrelated sources – so called "virtual sources" – returned by the EVD of the correlation matrix, and demonstrates the inability of this approach to solve alone the BSS problem[15-16].

4. The principle of least spatial complexity

In order to find the missing unitary matrix **V**, the principle of least spatial complexity is introduced in this section. Inspired from the definition of the Shannon entropy in information theory[17], a novel concept, the spatial entropy H, is proposed. The authors suggest defining the spatial entropy H as

$$H = \sum_{i=1}^{n_s} \pi_i H_i = -\sum_{i=1}^{n_s} \pi_i \int_\Gamma P_i(\mathbf{r}) \ln P_i(\mathbf{r}) d\Gamma(\mathbf{r}) \,, \tag{13}$$

where H_i, $P_i(\mathbf{r})$ and π_i are the i-th spatial entropy element, the statistical spatial intensity and the averaged power of the i-th source, respectively. For the continuous case, upon defining the statistical spatial intensity $P_i(\mathbf{r})$ as

$$P_i(\mathbf{r}) = \frac{\mathbb{E}\left\{\left|s_i(\mathbf{r},\omega;\varpi)\right|^2\right\}}{\int_\Gamma \mathbb{E}\left\{\left|s_i(\mathbf{r},\omega;\varpi)\right|^2\right\} d\Gamma(\mathbf{r})} \,, \tag{14}$$

and the i-th averaged power π_i as

$$\pi_i = \int_\Gamma \mathbb{E}\left\{\left|s_i(\mathbf{r},\omega;\varpi)\right|^2\right\} d\Gamma(\mathbf{r}) \,, \tag{15}$$

the spatial entropy H reads

$$H = -\sum_{i=1}^{n_s} \int_\Gamma \mathbb{E}\left\{ \left| s_i(\mathbf{r}, \omega; \varpi) \right|^2 \right\} \ln \frac{\mathbb{E}\left\{ \left| s_i(\mathbf{r}, \omega; \varpi) \right|^2 \right\}}{\int_\Gamma \mathbb{E}\left\{ \left| s_i(\mathbf{r}, \omega; \varpi) \right|^2 \right\} d\Gamma(\mathbf{r})} d\Gamma(\mathbf{r}) . \quad (16)$$

In brief, the spatial entropy is constructed such that, the more complex the sound source distribution, the higher the value of H is and vice versa. This means that minimizing spatial entropy with respect to matrix **V** will favor the recovery of the sound sources with as simple a spatial structure as possible. In other words, the spatial entropy is a measure of the compactness of the sound source distribution. This corresponds very well to our physiological perception of a sound source.

Let us now rewrite the expression of H in an explicit way as a function of **V** in order to undertake its minimization. The i-th sound source for its l-th position vector can be represented as

$$s_i(\mathbf{r}_l, \omega; \varpi) = \mathbf{e}_l^T \mathbf{\Phi} \mathbf{A} \mathbf{E}_i \mathbb{S} , \quad (17)$$

where $\mathbf{e}_l$ is the l-th column of the identity matrix and $\mathbf{E}_i = \mathbf{e}_i \mathbf{e}_i^T$ is a matrix whose all elements are zero except the i-th diagonal entry which equals one (superscript '^{T}' stands for the transposition operation). Therefore $s_i(\mathbf{r}_l, \omega; \varpi)$ has the statistical spatial intensity

$$P_i(\mathbf{r}_l) = \frac{\mathbf{e}_l^T \mathbf{\Phi} \mathbf{A} \mathbf{E}_i \mathbb{E}\left\{ \mathbb{S}\mathbb{S}^H \right\} \mathbf{E}_i \mathbf{A}^H \mathbf{\Phi}^H \mathbf{e}_l}{\sum_{l=1}^{N} \mathbf{e}_l^T \mathbf{\Phi} \mathbf{A} \mathbf{E}_i \mathbb{E}\left\{ \mathbb{S}\mathbb{S}^H \right\} \mathbf{E}_i \mathbf{A}^H \mathbf{\Phi}^H \mathbf{e}_l} = \frac{\mathbf{e}_l^T \mathbf{\Phi} \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \mathbf{\Phi}^H \mathbf{e}_l}{tr\left\{ \mathbf{\Phi} \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \mathbf{\Phi}^H \right\}} , \quad (18)$$

with $\mathbf{v}_i$, the i-th row of matrix **V** and tr, the trace of matrix $\mathbf{\Phi} \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \mathbf{\Phi}^H$. After substituting Eq. (18) into Eq. (16) and discretizing the sound source distribution as in Section 2, the spatial entropy reads

$$H(\mathbf{V}) = -\sum_{i=1}^{n_s} \sum_{l=1}^{N} (\mathbf{e}_l^T \mathbf{\Phi} \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \mathbf{\Phi}^H \mathbf{e}_l) \ln \left(\frac{\mathbf{e}_l^T \mathbf{\Phi} \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \mathbf{\Phi}^H \mathbf{e}_l}{tr\left\{ \mathbf{\Phi} \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \mathbf{\Phi}^H \right\}} \right) . \quad (19)$$

This cost function H(**V**) is now to be minimized with respect to matrix **V** under the constraint of unitarity of the latter. A very elegant approach to this difficult problem is to explicitly perform the optimization on the Stiefel manifold (ensemble of all unitary matrices). A recently proposed conjugate gradient (CG) algorithm owning outstanding optimization performance is used here, yet not detailed due to lack of place – (see Refs. [18-19]).

5. Experiment

5.1 Experimental setup

To assess the separation performance of the proposed algorithm, an experiment was conducted in a semi-anechoic room. The experimental setup is depicted in Fig. 2. A 60-microphone array is placed parallel to loudspeaker membranes which produce the sound sources. The distance between the microphone array and the loudspeakers is represented by z. The spacing between the centers of the loudspeaker membranes is defined by D. Different parameter combinations (i.e. z, D and the number n_s of the loudspeakers) have been investigated respectively. Near-field as well as far-field configurations are investigated. Loudspeakers are fed by mutually uncorrelated white noises with different power levels. All signals are sampled at 16384 Hz. The snapshot length is set to 4s, with 50% overlap and use of a Hanning window; the total number of the snapshot is 119.

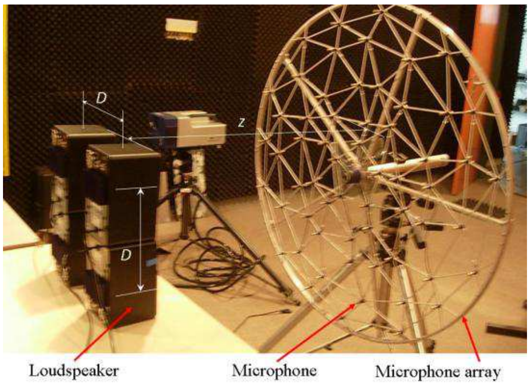

Figure 2. Experimental setup for blind separation of uncorrelated sound sources

5.2 Experimental results

The recordings from the microphones were then processed with the proposed algorithm. The global source distribution in the plane of the loudspeaker membranes was first reconstructed using the approach described in Section 2. The distribution was then decomposed into several uncorrelated sources by using the principle of least spatial complexity described in Section 4. Separation results for the case 'z=100cm, D=18cm and n_s =4' are illustrated in Fig. 3.

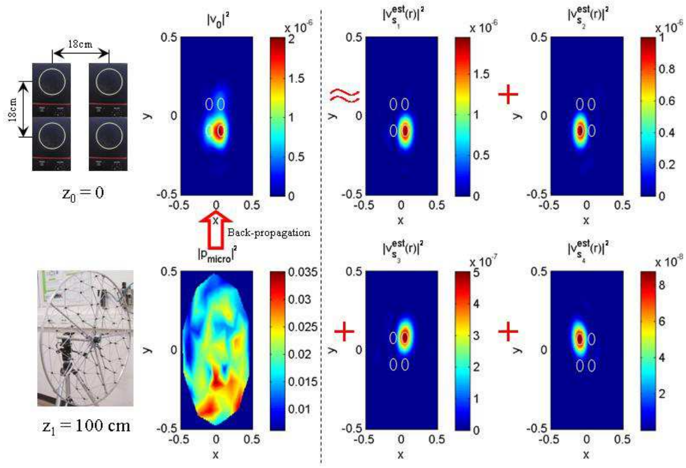

Figure 3. Separated sound sources at 1800Hz

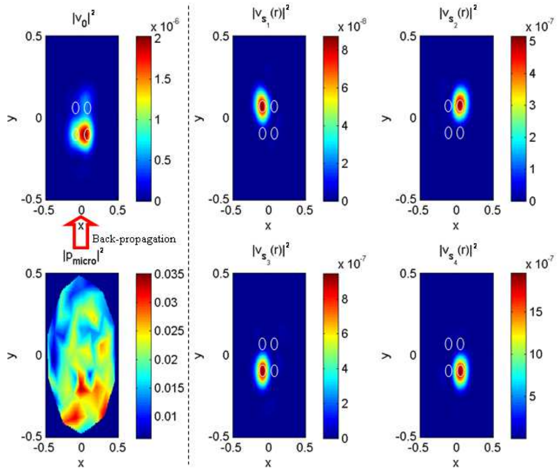

Figure 4. References sound sources at 1800Hz

The first upper left panel in Fig. 3 shows the configurations of the four loudspeakers; the lower left panel shows the irregular 60-microphone array; '$|p_{micro}|^2$' is the average power of the acoustic pressure interpolated in the plane of the microphone array; '$|V_0|^2$' is the average power of the backpropagated normal component of particle velocity in the plane of the loudspeaker membranes and the other four panels are the average power of normal component of particle velocity corresponding to the four separated sound sources. (Note that the final separation results are all displayed in terms of the particle velocity because its spatial resolution is usually much better than for the acoustic pressure.) It is seen that the globally backpropagated source distribution can not distinguish the existence of four sound sources: there is basically only one "hot spot" positioned on the loudest source. On the other hand, the proposed BSS algorithm perfectly singles out the four different sources which are perfectly localized on the four loudspeaker membranes.

In order to evaluate the separation performance of the proposed algorithm, a series of additional experiments were conducted in which only one loudspeaker was switched on each time. Four groups of measurements corresponding to four "reference sound sources" were independently processed (back-propagation from the microphone array to the plane of the sound sources). All reconstructed source distributions are depicted in Fig. 4. The four pictures on the right side of the dashed line display the average power of normal component of particle velocity of the four sources. Comparison of Figs. 3 and 4, it evidences a perfect match between the blind separation results and the controlled ones, both in terms of localization and of magnitude (even for the sound source with the smallest power magnitude). The only difference is in the order of the separated sound sources, which obviously must be arbitrary. Indeed, this fundamental indeterminacy inherent to all BSS algorithms is not at all a practical limitation.

5.3 Discussion

A large number of similar experiments with various parameter combinations (i.e. z, D and n_s) demonstrated the validity of the proposed decomposition algorithm in a wide frequency range (compatible with the backpropagation requirements). Three important aspects are further discussed based on these results, namely the effect of distance z between the sources and the array, of spacing D between the sources, and the estimation of the number n_s of sources.

As described in Refs. [20] and [14], the distance z between the loudspeakers and the microphone array has a direct impact on the expected reconstruction's quality, especially in terms of magnitude. In this regards, blind separation will suffer from exactly the same limitations as the global backpropagation (i.e. the reconstruction errors on the global source field $s(\mathbf{r},\omega;\varpi)$ will be inherited by its separated components $s_i(\mathbf{r},\omega;\varpi)$).

As for the spacing D between the centers of the loudspeaker membranes, it determines the spatial resolution limit of the algorithm. For instance, global backpropagation was unable to resolve between the loudspeakers with the settings in Fig. 3. One remarkable benefit of the source separation is that it is able to single out spatially overlapping sources, even below the spatial resolution limit. This is of considerable importance for source identification. In contrast, it means all sound sources could be overlapped together so that they would be regarded as a whole and large source, when D tends to be an enough small threshold, in which case source separation will no longer make sense.

In the previous example, the number n_s of sound sources to be separated was set to be equal to the actual number, say n_a, of loudspeakers. Further experiments have shown that when n_s is smaller than n_a, the separated sound sources usually correspond to the actual most powerful ones as a result of selecting the greatest eigenvalues in the whitening step (see Section 3). On the other hand, when n_s is greater than n_a, the extra separated sources mainly account for additive noise. This means the proposed algorithm has ability to correctly estimate the actual number of the uncorrelated sound sources, provided eigenvalues of the signal subspace are greater than those of the noise subspace. By way of an example, the blind separation of four sound sources (n_a =4) at 1800Hz, is illustrated in Figs. 5 and 6 with estimated source number n_s set to 3 and 6, respectively. Figure 5 exemplifies a good reconstruction of the three most powerful sources, despite the fourth one is not accounted for in the separation. In Fig. 6, although the two extra separated sources look very much like compact sources, their power magnitudes are orders of magnitude smaller than the other sound sources and thus can actually be categorized as noise.

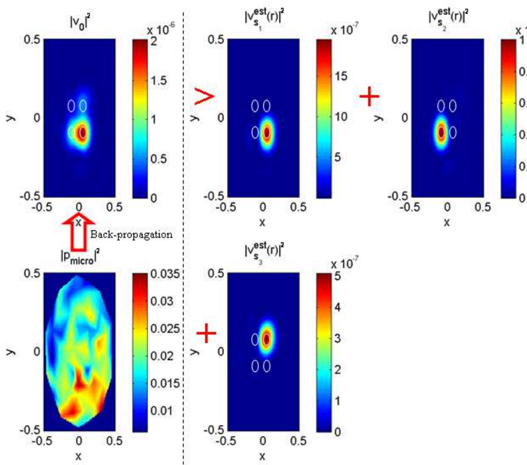

Figure 5. Separated sound sources with n_s =3 at 1800Hz

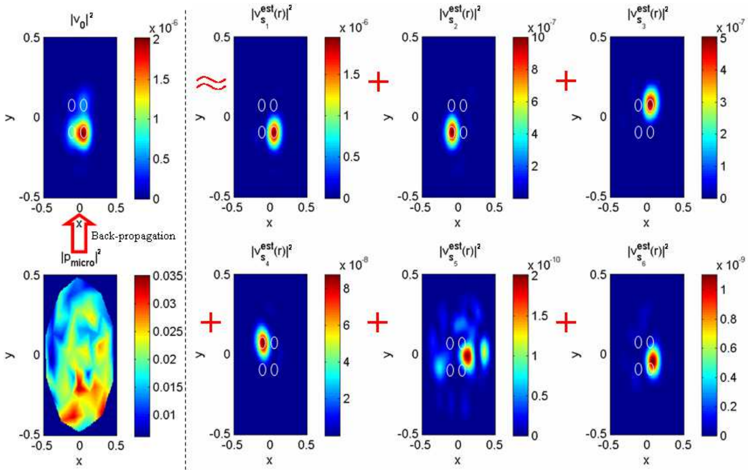

Figure 6. Separated sound sources with n_s =6 at 1800Hz

6. Conclusion

The proposed BSS algorithm is based on two simple assumptions. The first one is the mutual decorrelation of the sound sources of interest and the second one is their spatial compactness, or "least spatial complexity". It is believed that these two assumptions may be easily met in many industrial applications. It has been shown that a unique separation of sources is then possible by minimizing a spatial entropic cost function.

Experimental results have demonstrated excellent performances of the algorithm, and in particular its ability to go much beyond the spatial resolution limit allowed by standard backpropagation. The algorithm also has the ability to find the exact number of actual sound sources by inspecting their relative powers.

It must be emphasized, however, that the method is not tailored to decompose correlated sound sources, as would be the case with reflections for instance. Yet it is believed that the proposed principle of least spatial complexity would still achieve separation in this situation, an objective which is left for future work.

REFERENCES

1. B. Lafon, J. Antoni, M. Sidahmed and L. Polac. The Concept of Cyclic Sound Intensity and its Application to Acoustical Imaging, *International Journal of Sound and Vibration*, **330** (9), 2107–2121, (2011).
2. Pierre Comon. Independent Component Analysis: A New Concept?, *International Journal of Signal Processing*, **36** (3), 287–314, (1994).
3. Jean-François Cardoso. Blind Signal Separation: Statistical Principles, *International Journal of Proceeding of the IEEE*, **86** (10), 2009–2025, (1998).
4. Anthony J. Bell and Terrence J. Sejnowski. An Information-Maximisation Approach to Blind Separation and Blind Deconvolution, *International Journal of Neural Computation*, **7** (6), 1004–1034, (1995).
5. Te-Won Lee, M. Girolami, A. J. Bell and T. J. Sejnowski. A Unifying Information-Theoretic Framework for Independent Component Analysis, *International Journal of Computers & Mathematics with Applications*, **39**, 1–21, (2000).
6. Xi-Ren Cao and Ruey-wen Liu. General Approach to Blind Source Separation, *International Journal of IEEE Transactions on Signal Processing*, **44** (3), 562–571, (1996).
7. Hyvärinen, E. Oja . Independent Component Analysis: Algorithms and Application, *International Journal of Neural Networks*, **13** (4–5), 411–430, (2000).
8. Hyvärinen, J. Karhunen, E. Oja. *Independent Component Analysis*, Wiley, New York, 147–163, (2001).
9. Kennech E. Hild II, Deniz Erdogmus and Jose C. Principe. An Analysis of Entropy Estimators for Blind Source Separation, *International Journal of Signal Processing*, **86**, 181–194, (2006).
10. Liangfeng Guo and Marc Garland. The Use of Entropy Minimization for the Solution of Blind Source Separation Problems in Image Analysis, *International Journal of Pattern Recognition*, **39**, 1066–1073, (2006).
11. Sarit Shwartz, Yoav Y. Schechner and Michael Zibulevsky. Blind Separation of Convolutive Image Mixtures, *International Journal of Neurocomputing*, **71**, 2164–2179, (2008).
12. Xi-Lin Li and Tülay Adali. Blind Separation of Noncircular Correlated Sources Using Gaussian Entropy Rate, *International Journal of IEEE Transactions on Signal Processing*, **59** (6), 2969–2975, (2011).
13. Dinh Tuan Pham. Blind Separation of Instantaneous Mixture of Sources via the Gaussian Mutual Information Criterion, *International Journal of Signal Processing*, **81** (4), 855–870, (2001).
14. J. Antoni. Focalisation Bayésienne: Une Approche Unifiée du Problème Inverse en Acoustique, *CFA2010, 10ème Congrès Français d'Acoustique*, Lyon, France, 12–16 Avril, (2010).
15. Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space, *International Journal of Philosophical Magazine*, **2** (6), 559–572, (1901).
16. Jonathon Shlens. A Tutorial on Principal Component Analysis, Version 3.01, 22 April, (2009).
17. E. Shannon. A Mathematical Theory of Communication, *the Bell System Technical Journal*, **27**, 379–423, 623–656, (1948).
18. T. Abrudan, J. Eriksson, V. Koivunen. Conjugate Gradient Algorithm for Optimization under Unitary Matrix Constraint, *International Journal of Signal Processing*, **89** (9), 1704–1714, (2009).
19. Traian Abrudan. *Advanced Optimization Algorithms for Sensor Arrays and Multi-Antenna Communications*, Doctoral Dissertation, Department of Signal Processing and Acoustics, Helsinki University of Technology, Espoo, 33–50, (2008).
20. E. G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, 1–293, (1999).