BLIND SOURCE SEPARATION IN THE PRESENCE OF DOPPLER FREQUENCY SHIFTS

Arie Yeredor

School of Electrical Engineering
Tel-Aviv University, Tel-Aviv, 69978, ISRAEL
e-mail: `arie@eng.tau.ac.il`

ABSTRACT

We address the problem of blind separation of sources, mixed subject to possible Doppler frequency-shifts, differing between sources and between sensors. This situation is likely to occur, e.g., in scenarios involving mobile sensors and / or sources, but so far the multiple sources case does not seem to have been addressed (at least not in open literature, to our knowledge). We propose a batch-type iterative procedure for the estimation of the (static) mixing parameters and the frequency-shifts, followed by application of the inverse system to reconstruct the sources. The estimation procedure can be regarded as parameterized joint diagonalization of a bulk of rank-one matrices. Somewhat surprisingly, correlation matrices at zero lag are generally sufficient for separation, and consequently, the separation of white Gaussian sources is generally possible in this framework - unlike the situation in classical static mixing - as we demonstrate in simulation.

1. INTRODUCTION AND PROBLEM FORMULATION

The basic, classical Blind Source Separation (BSS) problem (often also termed Independent Components Analysis (ICA)) addresses the separation of statistically independent sources from noiseless observations of their static linear mixture. A vast variety of expansions of the basic model (e.g., noisy / convolutive / nonlinear mixtures, etc.) have been studied extensively in the past decade. However, although likely in context of mobile sensors and / or sources, so far the framework of Doppler-contaminated BSS has not (to our knowledge) been considered in the (open) literature. The basic noiseless model accounting for the effect (after conversion to baseband) is the following:

$$x_k[n] = \sum_{\ell=1}^{L} a_{k\ell} e^{j\omega_{k\ell} n} s_\ell[n] \quad , \quad k = 1, 2, ..., K \tag{1}$$

where $s_1[n], s_2[n], ...s_L[n]$ are independent sources (complex-valued, after conversion to baseband) and $x_1[n], x_2[n], ...x_K[n]$ are the observations. We assume that $K \geq L$. The set $\{a_{k\ell}\}$ are unknown (complex-valued) linear mixture coefficients (gains / attenuations and phase-shifts) and $\{\omega_{k\ell}\}$ are unknown Doppler frequency-shifts (from source ℓ to sensor k). Since the "absolute" frequencies of the sources are unknown (and, actually, are often undefined), we may assume arbitrarily, without loss of generality, that an unshifted version appears, e.g., at the first sensor, namely that $w_{1\ell} = 0$ for all $\ell = 1, 2, ...L$.

Note that the model (1) does not account for possible time-delay differences between sensors, which may occur when the delays associated with the spatial aperture of the receivers' array are not negligible with respect to the correlation-lengths of the sources. Indeed, for mobile sensors, relatively large apertures may be expected. Likewise, if mobile sources are involved, usually only large apertures would inflict different Doppler shifts between sensors. However, it is important to realize, that there can always be sources that are narrow-banded enough on one hand, but whose central frequencies are high enough on the other hand - such that the Doppler effect is dominant and the delay differences are negligible, no matter how large the aperture is. In this work we chose to address only the Doppler effect, thus giving rise to (1). The complementary case of BSS with pure delay differences (and no Doppler effect) has been considered, e.g., in [1, 2]. The combination of both is a much more complex (yet practical) problem, and one of the purposes of this work is to establish a basis for solving the "pure Doppler" case first, as additional foundations for continued work on the combined Doppler-delays BSS problem.

Observe that the model (1) can also be written in matrix-form as

$$\boldsymbol{x}[n] = (\boldsymbol{A} \odot \boldsymbol{D}[n; \boldsymbol{\Omega}])\, \boldsymbol{s}[n] \tag{2}$$

where $\odot$ denotes Hadamard's (element-wise) matrix multiplication, $\boldsymbol{A}$ denotes the $K \times L$ matrix of the linear mixing coefficients $\{a_{kl}\}$ and $\boldsymbol{D}[n;\boldsymbol{\Omega}]$ is the Doppler-shifts matrix, whose (k,ℓ)-th element is given by $\boldsymbol{D}_{k\ell}[n;\boldsymbol{\Omega}] = e^{j\omega_{k\ell}n}$, with the $K \times L$ matrix $\boldsymbol{\Omega}$ encompassing all the frequency-shifts $\{\omega_{k\ell}\}$.

As will be shown later, the problem can often be resolved using second-order statistics alone. More specifically, assuming zero-mean sources, we shall only exploit the fact that no pair of sources is correlated at zero-lag. Additionally, in order to mitigate the scale-ambiguity (inherent in all BSS problems), we shall employ the "working assumption" that all sources have unit variance (see, e.g., [3] for justification). These assumptions are summarized by

$$E\left[\boldsymbol{s}[n]\boldsymbol{s}^H[n]\right] = \boldsymbol{I}. \tag{3}$$

Note that only power-stationarity of the sources is assumed. In general, these assumptions are quite modest comparing to the "classical" BSS assumptions, since it turns out (somewhat surprisingly) that thanks to the presence of the Doppler frequency shifts, there's no need for an *additional* "diagonal property". Such an additional diagonal property is required in the basic BSS problem (see, e.g., [4]), and is usually satisfied by the sources' independence (e.g., through their joint cumulants matrices [5]) or wide-sense stationarity (e.g., through their joint correlation matrices at different time-lags [6]). An intriguing implication of the elimination of this requirement is the ability to separate (in the presence of Doppler-shifts) spectrally-white Gaussian sources - which are not separable in the classical basic problem [3]. We shall demonstrate this ability using simulation results in section 3.

2. DERIVATION OF THE ESTIMATION ALGORITHM

Our proposed separation approach consists of two parts: estimation of the mixing parameters (coefficients and Doppler shifts), followed by application of the inverse system to the observations. Thus, we work in batch-mode, and assume that N observations of $\boldsymbol{x}[n]$, $n = 1, 2, ..., N$ are available.

Although the sources may be stationary, the observations vector is definitely non-stationary in the presence of Doppler-shifts, as can be regarded from

$$E\left[\boldsymbol{x}[n]\boldsymbol{x}^H[n]\right] = \left(\boldsymbol{A} \odot \boldsymbol{D}[n;\boldsymbol{\Omega}]\right)\left(\boldsymbol{A} \odot \boldsymbol{D}[n;\boldsymbol{\Omega}]\right)^H. \tag{4}$$

Consequently, standard straightforward time-averaging for estimation of the observations' statistical properties may prove futile. We therefore choose to resort to an approach involving the concept of parameterized joint diagonalization of a bulk of matrices. To this end, observe that at each time-instant n, the product $\boldsymbol{x}[n]\boldsymbol{x}^H[n]$ can be viewed as a noisy realization of the true correlation matrix at that time-instant, namely

$$\boldsymbol{x}[n]\boldsymbol{x}^H[n] \approx \left(\boldsymbol{A} \odot \boldsymbol{D}[n;\boldsymbol{\Omega}]\right)\left(\boldsymbol{A} \odot \boldsymbol{D}[n;\boldsymbol{\Omega}]\right)^H. \tag{5}$$

A least-squares (LS) based estimation strategy would then be to minimize the associated "error" norm:

$$\min_{\boldsymbol{A},\boldsymbol{\Omega}} \sum_{n=1}^{N} \left\| \boldsymbol{x}[n]\boldsymbol{x}^H[n] - \left(\boldsymbol{A} \odot \boldsymbol{D}[n;\boldsymbol{\Omega}]\right)\left(\boldsymbol{A} \odot \boldsymbol{D}[n;\boldsymbol{\Omega}]\right)^H \right\|_F^2$$

where $\|\cdot\|_F^2$ denotes the squared Frobenius norm[1]. Our minimization approach will be based on extending and modifying an existing joint-diagonalization algorithm [7] (see also [2]).

Let us denote by $\boldsymbol{a}_\ell$ and $\boldsymbol{d}_\ell[n]$ (respectively) the ℓ-th columns of $\boldsymbol{A}$ and $\boldsymbol{D}[n;\boldsymbol{\Omega}]$ (omitting $\boldsymbol{\Omega}$ from $\boldsymbol{d}_\ell[n]$ for notational convenience). We wish to minimize

$$\sum_{n=1}^{N} \left\| \boldsymbol{x}[n]\boldsymbol{x}^H[n] - \sum_{\ell=1}^{L} \left(\boldsymbol{a}_\ell \odot \boldsymbol{d}_\ell[n]\right)\left(\boldsymbol{a}_\ell \odot \boldsymbol{d}_\ell[n]\right)^H \right\|_F^2.$$

Note that the vector $\boldsymbol{a}_\ell \odot \boldsymbol{d}_\ell[n]$ can also be written as $\breve{\boldsymbol{D}}^{(\ell)}[n]\boldsymbol{a}_\ell$, where $\breve{\boldsymbol{D}}^{(\ell)}[n]$ is a diagonal (and unitary) matrix whose diagonal is comprised of the elements of $\boldsymbol{d}_\ell[n]$, namely $\breve{\boldsymbol{D}}^{(\ell)}_{mm} = \left(\boldsymbol{d}_\ell[n]\right)_m = e^{j\omega_{m\ell}n}$. Using this notation, we shall now derive the minimization of the LS criterion with respect to (w.r.t.) the ℓ-th column of $\boldsymbol{A}$ ($\boldsymbol{a}_\ell$), treating all other parameters as constants. We shall then proceed to derive the minimization w.r.t. the ℓ-th column of $\boldsymbol{\Omega}$. The final algorithm will consist of "sweeps" alternating between L minimization operations w.r.t. all columns of $\boldsymbol{A}$ and L minimization operations w.r.t. all columns of $\boldsymbol{\Omega}$. Monotonic decrease of the LS criterion will thus be guaranteed.

Following estimation of $\boldsymbol{A}$ and $\boldsymbol{\Omega}$, straightforward estimates of the sources are obtained from $\hat{\boldsymbol{s}}[n] = \left(\widehat{\boldsymbol{A}} \odot \boldsymbol{D}[n,\widehat{\boldsymbol{\Omega}}]\right)^\dagger \boldsymbol{x}[n]$, where $\widehat{\boldsymbol{A}}$ and $\widehat{\boldsymbol{\Omega}}$ denote the estimates of $\boldsymbol{A}$ and $\boldsymbol{\Omega}$, respectively, and the † superscript denotes the (pseudo-)inverse.

2.1. Minimization w.r.t. the mixing coefficients (a column of $\boldsymbol{A}$)

The LS criterion under minimization can be written as

$$C(\boldsymbol{a}_\ell) \triangleq \sum_{n=1}^{N} \left\| \boldsymbol{P}^{(\ell)}[n] - \breve{\boldsymbol{D}}^{(\ell)}[n]\boldsymbol{a}_\ell\boldsymbol{a}_\ell^H\breve{\boldsymbol{D}}^{(\ell)H}[n] \right\|_F^2 \tag{6}$$

[1] The Frobenius norm of a matrix $\boldsymbol{Q}$ is given, e.g., by $\|\boldsymbol{Q}\|_F^2 = \mathrm{Trace}\{\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{H}}\}$.

where

$$\boldsymbol{P}^{(\ell)}[n] \triangleq \boldsymbol{x}[n]\boldsymbol{x}^H[n] - \sum_{\substack{m=1\\m\neq\ell}}^{L} \check{\boldsymbol{D}}^{(m)}[n]\boldsymbol{a}_m\boldsymbol{a}_m^H\check{\boldsymbol{D}}^{(m)H}[n].$$

Due to the unitarity of $\check{\boldsymbol{D}}^{(\ell)}[n]$ (and the invariance of the Frobenius norm under unitary transformations), $C(\boldsymbol{a}_\ell)$ can be rewritten as

$$C(\boldsymbol{a}_\ell) = \sum_{n=1}^{N} \left\| \check{\boldsymbol{D}}^{(\ell)}[n]\boldsymbol{P}^{(\ell)}[n]\check{\boldsymbol{D}}^{(\ell)H} - \boldsymbol{a}_\ell\boldsymbol{a}_\ell^H \right\|_F^2. \quad (7)$$

Defining, for convenience, the (Hermitian) matrix

$$\widetilde{\boldsymbol{P}}^{(\ell)}[n] \triangleq \check{\boldsymbol{D}}^{(\ell)}[n]\boldsymbol{P}^{(\ell)}[n]\check{\boldsymbol{D}}^{(\ell)H}, \quad (8)$$

$C(\boldsymbol{a}_\ell)$ can be expressed as

$$\begin{aligned} C(\boldsymbol{a}_\ell) &= \sum_{n=1}^{N} \mathrm{Trace}\left\{ (\widetilde{\boldsymbol{P}}^{(\ell)}[\mathrm{n}] - \boldsymbol{a}_\ell\boldsymbol{a}_\ell^{\mathrm{H}})(\widetilde{\boldsymbol{P}}^{(\ell)}[\mathrm{n}] - \boldsymbol{a}_\ell\boldsymbol{a}_\ell^{\mathrm{H}}) \right\} \\ &= \sum_{n=1}^{N} \left[\mathrm{Trace}\left\{ \widetilde{\boldsymbol{P}}^{(\ell)}[\mathrm{n}]\widetilde{\boldsymbol{P}}^{(\ell)}[\mathrm{n}] \right\} - 2\boldsymbol{a}_\ell^{\mathrm{H}}\widetilde{\boldsymbol{P}}^{(\ell)}[\mathrm{n}]\boldsymbol{a}_\ell + \left(\boldsymbol{a}_\ell\boldsymbol{a}_\ell^{\mathrm{H}}\right)^2 \right] \\ &= C' - 2\boldsymbol{a}_\ell^H \left(\sum_{n=1}^{N} \widetilde{\boldsymbol{P}}^{(\ell)}[n] \right) \boldsymbol{a}_\ell + N\left(\boldsymbol{a}_\ell^H\boldsymbol{a}_\ell\right)^2, \quad (9) \end{aligned}$$

where C' is some constant, irrelevant for minimization w.r.t. $\boldsymbol{a}_\ell$. To proceed, we decompose $\boldsymbol{a}_\ell$ as $\boldsymbol{a}_\ell \triangleq a\boldsymbol{\alpha}$, where a is some real-valued constant and $\boldsymbol{\alpha}$ is a unitary vector, $\boldsymbol{\alpha}^H\boldsymbol{\alpha} = 1$. Consequently, we now need to minimize w.r.t. a and $\boldsymbol{\alpha}$,

$$C(\boldsymbol{a}_\ell) = Na^4 - 2a^2\boldsymbol{\alpha}^H \left(\sum_{n=1}^{N} \widetilde{\boldsymbol{P}}^{(\ell)}[n] \right) \boldsymbol{\alpha} + C' \quad (10)$$

Obviously, minimization w.r.t. $\boldsymbol{\alpha}$ (under its unitarity constraint) requires $\boldsymbol{\alpha}$ to be the eigenvector $\boldsymbol{v}_1$ associated with the maximum eigenvalue λ_1 of the matrix $\frac{1}{N}\sum_{n=1}^{N} \widetilde{\boldsymbol{P}}^{(\ell)}[n]$. Substituting back into (10) requires to minimize $a^4 - 2a^2\lambda_1$ w.r.t. a. When λ_1 is negative, the minimizing a is zero, and consequently in such cases the minimizing $\boldsymbol{a}_\ell$ is $\boldsymbol{0}$. However, usually λ_1 is positive, in which case the minimizing solution is $a = \sqrt{\lambda_1}$ and $\boldsymbol{a}_\ell = \sqrt{\lambda_1} \cdot \boldsymbol{v}_1$. Note that the complex phase of $\boldsymbol{v}_1$ (hence of $\boldsymbol{a}_\ell$) is irrelevant for the minimization. This is in accordance with the residual phase ambiguity inherent in complex-valued BSS problems, since the complex phase of the mixing coefficients can be commuted with a constant phase-shift of the sources.

2.2. Minimization w.r.t. the frequency shifts (a column of $\boldsymbol{\Omega}$)

We now proceed to minimize the LS criterion w.r.t. a single column of $\boldsymbol{\Omega}$, treating all the other columns, as well as all the mixing coefficients $\boldsymbol{A}$, as constants. Observing (9), it is evident that with $\boldsymbol{a}_\ell$ constant, minimization of the criterion requires maximization of $\boldsymbol{a}_\ell^H \left(\sum_{n=1}^{N} \widetilde{\boldsymbol{P}}^{(\ell)}[n] \right) \boldsymbol{a}_\ell$, or, more explicitly, reinstating the frequency-shifts into the expression (8) for $\widetilde{\boldsymbol{P}}^{(\ell)}[n]$, we wish to minimize

$$\sum_{p=1}^{K}\sum_{q=1}^{K} a_{p\ell}^* \left[\sum_{n=1}^{N} \boldsymbol{P}_{pq}^{(\ell)}[n] e^{j(\omega_{p\ell}-\omega_{q\ell})n} \right] a_{q\ell} \quad (11)$$

w.r.t. $\{\omega_{p\ell}\}_{p=2}^{K}$ (all frequencies in the ℓ-th column, except for $\omega_{1\ell}$, which has been arbitrarily set to zero). In general, the minimization of (11) w.r.t. the entire set requires the calculation of the discrete-time Fourier transforms (DTFTs) of the sequences $\left\{\boldsymbol{P}_{pq}^{(\ell)}[n]\right\}$, followed by a relatively simple $(K-1)$-dimensional search. We shall not describe this procedure in detail in here, due to lack of space. Instead, we shall only explore the (relatively) simple $K = L = 2$ case: Observe that in this case, only two elements (out of four) in the outer summation in (11) depend on the unknown frequencies - the elements corresponding to $(p, q) = (1, 2)$ or $(2, 1)$. Thus, for $\ell = 1$ we seek to maximize, w.r.t. ω_{21}:

$$a_{11}^* a_{21} \sum_{n=1}^{N} \boldsymbol{P}_{12}^{(1)}[n] e^{-j\omega_{21}n} + a_{21}^* a_{11} \sum_{n=1}^{N} \boldsymbol{P}_{21}^{(1)}[n] e^{j\omega_{21}n}.$$

Observing that due to the conjugate-symmetric structure of $\boldsymbol{P}^{(\ell)}[n]$ these two terms are a conjugate pair, we end up maximizing

$$\max_{\omega} \mathrm{Real}\left\{ a_{11}^* a_{21} q_1(e^{j\omega}) \right\} \quad (12)$$

where $q_1(\mathrm{e}^{j\omega}) \triangleq \sum_{n=1}^{N} \boldsymbol{P}_{12}^{(1)}[n]e^{-j\omega n}$ is the DTFT of the sequence $\left\{\boldsymbol{P}_{12}^{(1)}[n]\right\}$, which may be searched for the maximizing ω in a pre-defined feasibility-range where the expected Doppler shifts can exist. Computationally-efficient techniques, such as the zoom-FFT or the chirp-FFT (e.g., [8]) can be used for the computation of the DTFT within such a range.

Likewise, for $\ell = 2$ (seeking ω_{22}), we need to maximize $\mathrm{Real}\left\{ a_{12}^* a_{22} q_2(e^{j\omega}) \right\}$ with $q_2(\mathrm{e}^{j\omega}) \triangleq \sum_{n=1}^{N} \boldsymbol{P}_{12}^{(2)}[n]e^{-j\omega n}$.

2.3. An initial guess

For such alternating-directions type algorithms, an "intelligent" initial guess is often required, so as to avoid convergence to spurious minima. Luckily, a reasonable guess for the Doppler shifts can be readily obtained from the data, as follows. Observe the sequence

$$r_k[n] \stackrel{\triangle}{=} x_k[n]x_1^*[n] = \sum_{\ell=1}^{L}\sum_{m=1}^{L} a_{k\ell}a_{1m}^* e^{j(\omega_{k\ell}-\omega_{1m})n} s_\ell[n]s_m^*[n].$$

Recalling that by convention $\omega_{1m} = 0 \quad \forall m$, and that $E\left[s_\ell[n]s_m^*[n]\right] = \delta_{\ell m}$ (Kronecker's delta), we may rewrite $r_k[n]$ as:

$$r_k[n] = \sum_{\ell=1}^{L} a_{k\ell}a_{1\ell}^* e^{j\omega_{k\ell}n} + \sum_{\ell=1}^{L}\sum_{m=1}^{L} a_{k\ell}a_{1m}^* v_{\ell m}[n]e^{j\omega_{k\ell}n} \tag{13}$$

where $v_{\ell m}[n]$ are zero-mean, mutually uncorrelated "noise" signals. In the common case where all the sources are spectrally white, all $v_{\ell m}[n]$ are spectrally white as well. Thus, while the first term in (13) is the superposition of L tones, each at the unknown frequency $\omega_{k\ell}$, the second term is often a white noise term. Consequently, under reasonable conditions on the frequencies and on the mixing coefficients, estimates of the set of frequencies $\omega_{k\ell}$ (for $\ell = 1, 2, ..., L$) can be readily extracted from the magnitude of the DTFT of the sequence $r_k[n]$ as obvious spectral peaks. Note, however, that since these peaks are not labeled, their consistent association with the sources (ℓ) can be ambiguous for $K > 2$. Nevertheless, in such cases the differences between frequencies can be extracted in a similar way, and help resolve these ambiguities (yet, the problem becomes quite complex for K, L much larger than 2).

As for the mixing coefficients - we currently do not have a straightforward procedure for an intelligent initial guess in the general case. In the $K = L = 2$ case, a (nonlinear) algebraic scheme can be used for obtaining reasonable estimates, but this scheme will not be pursued in here. In our simulation we used the identity matrix as an initial guess for $\boldsymbol{A}$.

3. SIMULATION RESULTS

We simulated a $K = L = 2$ case, with the mixing matrix $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ -1+j & 2-j \end{bmatrix}$ and the Doppler-shifts $\omega_{21} = -0.2$ and $\omega_{22} = 0.123$ (in radians). The two independent sources were white, zero-mean complex-Gaussian with unit variance. The observation length was $N = 1000$.

To illustrate the initial guess procedure, Figure (1) depicts the DTFT of the sequence $r_2[n]$ (taken from a single trial) in the range $\omega \in [-0.5, 0.5]$. The two peaks at $\omega \approx -0.2, 0.123$ are evident.

The performance was averaged over 1000 trials in terms of the Interference to Signal Ratio (ISR), computed (per trial) as follows. The time-varying

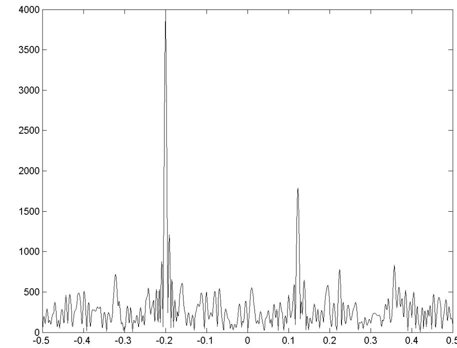

Fig. 1: Magnitude of the DTFT of $r_2[n]$

"contamination matrix" was computed as $\boldsymbol{T}[n] = \left(\widehat{\boldsymbol{A}} \odot \boldsymbol{D}[n, \widehat{\boldsymbol{\Omega}}]\right)^{-1} (\boldsymbol{A} \odot \boldsymbol{D}[n, \boldsymbol{\Omega}])$ and its element-wise absolute-squared sum was calculated over the observation interval, $\boldsymbol{T} = \sum_{n=1}^{N} \boldsymbol{T}[n] \odot \boldsymbol{T}^*[n]$. After resolving the permutation ambiguity, the ratios $\boldsymbol{T}_{12}/\boldsymbol{T}_{11}$ and $\boldsymbol{T}_{21}/\boldsymbol{T}_{22}$ were denoted ISR1 and ISR2, whose averaged values were -27.1dB & -29.3dB, respectively.

Similar performance was obtained under many other, but not under all mixing conditions. For example, the case $\omega_{21} = \omega_{22}$ is equivalent to a static mixture, with the Doppler-shift applied directly to $x_2[n]$ - obviously non-separable for Gaussian sources. The precise separability conditions and associated performance bounds remain to be explored.

4. REFERENCES

[1] B. Emile and P. Comon, "Estimation of time delays between unknown colored signals," *Signal Processing*, vol. 69, pp. 93–100, 1998.

[2] A. Yeredor, "Blind source separation with pure delays mixture," *Proceedings of ICA2001*, 2001.

[3] J.-F. Cardoso, "Blind signal separation: statistical principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, 1998.

[4] L. Parra and P. Sajda, "Blind source separation via generalized eigenvalue decomposition," *Journal of Machine Learning Research*, vol. 4, pp. 1261–1269, 2003.

[5] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non Gaussian signals," *IEE - Proceedings -F*, vol. 140, no. 6, pp. 362–370, 1993.

[6] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434–444, 1997.

[7] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation," *IEEE Trans. Signal Processing*, vol. 50, no. 7, pp. 1545–1553, 2002.

[8] A. V. Oppenheim, R. W. Schafer, and J Buck, *Discrete-time signal processing*, Prentice Hall, 2nd edition, 1999.