BLIND SOURCE DETECTION AND SEPARATION USING SECOND ORDER NON-STATIONARITY

Antoine Souloumiac

Schlumberger Industries, ETL/EMP
50, avenue Jean Jaurès, BP 620-05, 92542 Montrouge Cedex, France
souloum@montrouge.em.slb.com

ABSTRACT

We address the problem of using an array of sensors for detecting a narrow band source and separating its signal from unwanted disturbance signals, that is jammers and noise. The power of the desired signal is assumed to move from one level to another. This second order non-stationarity occurs, for instance, in frequency jumping systems, and more generally, at the beginning or at the end of any communication. We derive a method based on the generalized eigenstructure of two covariance matrices which requires no a priori knowledge of the array manifold, but only second order stationarity of the disturbance signals. The loss in signal to interference plus noise ratio (SINR) due to finite sample effect is calculated in closed form at the first order and validated by simulations. This last result shows that the method gives interesting performance in a wide range of situations.

1. INTRODUCTION

Let us denote M the number of sensors, $s(t)$ the narrow band signal emitted by the source of interest, and a the M-dimensional steering vector of this source. The contribution of any other sources in the frequency band of $s(t)$ is an M-dimensional signal denoted $n(t)$. In the following, s and n are assumed to be complex, zero mean, independent stochastic processes. The array output $x(t)$ is modeled as:

$$x(t) = a\, s(t) + n(t) \tag{1}$$

This model raises a classical issue of narrow band array processing: the separation of $s(t)$ from disturbance signal $n(t)$. We only consider here linear spatial filtering, which means that $\hat{s}(t)$, the estimation of $s(t)$, can be written as:

$$\hat{s}(t) = u^H x(t) \tag{2}$$

where u is an M-dimensional complex vector. Numerous estimators have been proposed, see for instance [1]. They use different kinds of a priori knowledge, like for example: a reference signal correlated with $s(t)$ and decorrelated with $n(t)$, the source steering vector a, or a parametrization of the antenna set of steering vectors, i.e. the array manifold. [1]

This paper is based on the assumption of:

- *second order non-stationarity* of the signal of interest s and
- *second order stationarity* of the disturbance signal n.

We assume that the power of $s(t)$, i.e. $E(|s(t)|^2)$, is equal to σ_1^2 in a first period of observation and equal to σ_2^2 in a second one. On the contrary, the second order stationarity of $n(t)$ implies that $R_n \overset{\text{def}}{=} E(n(t)n(t)^H)$ remains the same in both periods. Therefore, the covariance matrix of the array output can be written as:

$$R_1 \overset{\text{def}}{=} E(x(t)x(t)^H) = \sigma_1^2 aa^H + R_n \tag{3}$$

during the first period, and:

$$R_2 \overset{\text{def}}{=} E(x(t)x(t)^H) = \sigma_2^2 aa^H + R_n \tag{4}$$

during the second one.

This kind of second order non-stationarity happens for instance at the beginning and at the end of a communication. Besides, the modulation technique itself may cause a change of the source of interest power in a given frequency band, in frequency jumping systems for instance. To the best of our knowledge, this array processing problem has only been considered in [2]; a more complete comparison is done in section 3.

2. THE PROPOSED METHOD

The proposed method is based on the generalized eigenstructure of the couple (R_2, R_1). It is easy to check that

[1] In this last case, the source steering vector a is estimated via a direction of arrival estimation step.

(R_2, R_1) admits only two distinct generalized eigenvalues:

- 1 degenerated $M-1$ times and whose corresponding eigensubspace is the hyperplane orthogonal to a, and
- $\lambda \stackrel{\text{def}}{=} \mu_2/\mu_1$, where $\mu_2 \stackrel{\text{def}}{=} 1 + \sigma_2^2 a^H R_n^{-1} a$ and $\mu_1 \stackrel{\text{def}}{=} 1 + \sigma_1^2 a^H R_n^{-1} a$, corresponding to the eigenvector $v \stackrel{\text{def}}{=} R_n^{-1} a$. Note that this vector is the optimal linear spatial filter for estimating $s(t)$: the Spatial Matched Filter [1]. Without loss of generality, we can assume that $\sigma_1^2 \leq \sigma_2^2$, which implies that $\lambda \geq 1$.

Consequently, the proposed algorithm can be summarized as follows:

- Estimate the two covariance matrices $\widehat{R}_1$ and $\widehat{R}_2$,
- Compute their generalized eigenstructure,
- Detect if the power of a source has changed or not between the two periods of observation by testing the generalized eigenvalues,
- Estimate the signal of interest by filtering the array output with the corresponding generalized eigenvector $\hat{v}$, i.e. by computing:

$$\hat{s}(t) \stackrel{\text{def}}{=} \hat{v}^H x(t) \tag{5}$$

An interesting application of this algorithm consists of continuously watching a given narrow frequency band, detecting any source affected by a non-stationarity and filtering it in real time. This could be easily implemented by computing recursively the matrices $\widehat{R}_1$ and $\widehat{R}_2$ with two non-overlapping sliding windows. The delay between the two windows and their lengths are free parameters that can be adjusted to fit with the kind of non-stationarity to be detected and the required quality of separation (SINR). Note also that the computation of the eigenvector of maximal eigenvalue of the couple $(\widehat{R}_1, \widehat{R}_2)$ does not necessitate an actual generalized eigendecomposition. It is equivalent to the optimization:

$$\max_u \frac{u^H \widehat{R}_2 u}{u^H \widehat{R}_1 u} \tag{6}$$

Observe that the maximization of this Rayleigh quotient is easy to implement recursively.

The detection step can provide additional information. In particular, it allows to detect whether our modeling (1), (3) and (4) is valid or not. If the propagation of $s(t)$ is affected by multipaths, then the contribution to the source of interest is not necessarily monodimensional (like $a\,s(t)$) and (1) is not valid. If more than one source is affected by a non-stationarity between the first and the second period of observation, then (3) and (4) are not valid either. But, in both cases, the anomaly can be detected by considering the generalized eigenvalues of (R_2, R_1): in general, two or more eigenvalues are distinctly superior to 1. In the following, we no longer consider either this detection step, or the statistical test problem it contains. The next sections are devoted to the comparison of this technique with existing similar methods and to an asymptotical performance analysis of the estimation step.

3. COMPARISON WITH EXISTING METHODS

Considering two covariance matrices estimated before and after a second order non-stationarity of the signal of interest is not a new idea: it has already been proposed in [2]. But this method and our contribution do not necessitate the same assumptions and do not lead to the same algorithm.

In [2], the disturbance signal $n(t)$ is split up into the contribution of jammers and noise. The noise is assumed to be white and Gaussian and the number of jammers is assumed to be smaller than M. In addition, the signal $s(t)$ is assumed to be absent during the first period of information, i.e. $\sigma_1^2 = 0$. By comparison, our method needs less assumptions and is therefore more general.

The algorithm proposed in [2] can be summarized as follows:

- estimation of $\widehat{R}_1$ and $\widehat{R}_2$,
- estimation of the subspace spanned by the jammers steering vectors by means of the eigendecomposition of $\widehat{R}_1$,
- calculation of the spatial filter $\hat{v}$ by maximization of $u^H \widehat{R}_2 u / u^H u$ with u orthogonal to the jammers' subspace.

This algorithm necessitates two eigendecompositions instead of one generalized eigendecomposition in ours. In the second step, it solves the sometimes difficult problem of estimating the number of jammers; a problem which is avoided in our method. Furthermore, [2] converges to a spatial filter that cancels all the jammers, and it is well known that cancelling all the jammers leads to sub-optimal SINR, see for instance [1]. Finally, the asymptotical SINR reached by our estimator is particularly easy to evaluate theoretically when the numbers of samples used to estimate $\widehat{R}_1$ and $\widehat{R}_2$ are large but finite. This is the object of the following section.

4. ASYMPTOTIC PERFORMANCE ANALYSIS

4.1. Definitions and assumptions

A spatial filter $\hat{v}$ is estimated with the method described in section 2, then it is used to filter the array output $x(t)$:

$$\hat{s}(t) = \hat{v}^H x(t) = \hat{v}^H a\, s(t) + \hat{v}^H n(t) \tag{7}$$

The vector random variable $\hat{v}$ is assumed to be independent of $x(t)$, which means in particular that the spatial filter $\hat{v}$ is applied on data that have not been used to estimate it. In other words, $\hat{v}$ is used outside the two periods of observation. We define the mean SINR ρ provided by a given spatial filter $\hat{v}$, by:

$$\rho \stackrel{\text{def}}{=} \frac{\sigma^2 E(|\hat{v}^H a|^2)}{E(\hat{v}^H R_n \hat{v})} = \sigma^2 \frac{a^H E(\hat{v}\hat{v}^H) a}{\text{Tr}\ (R_n E(\hat{v}\hat{v}^H))} \tag{8}$$

where Tr means trace. Note that the SINR calculation reduces to the evaluation of the term $E(\hat{v}\hat{v}^H)$. Finally, we write ρ_{max} the maximal SINR obtained by linear filtering with any spatial filter proportional to v, i.e.:

$$\rho_{max} \stackrel{\text{def}}{=} \sigma^2 a^H R_n^{-1} a \tag{9}$$

where $\sigma^2 \stackrel{\text{def}}{=} E(|s(t)|^2)$

We assume that:

- n and s are Gaussian circular stochastic processes,
- $\widehat{R}_1$ and $\widehat{R}_2$ are estimated with, respectively, N_1 and N_2 independent samples of x,
- the two periods of observation do not overlap, which makes $\widehat{R}_1$ and $\widehat{R}_2$ independent.

4.2. Results and validation

After some calculations summarized in the appendix, we obtain at the first order in N_1^{-1} and N_2^{-1}:

$$1 - \frac{\rho}{\rho_{max}} = \mu_1\,(M-1)\left(\frac{1}{N_1} + \frac{1}{\lambda N_2}\right)\left(\frac{\lambda}{\lambda - 1}\right)^2 \tag{10}$$

A slightly more complicated formula can be found if n and s are not Gaussian. In this case, the formula involves the fourth order cumulants of n, but does not contain the fourth order cumulant of s. Consequently, (10) holds even if s is not Gaussian. This point deserves to be emphasized. It means that the mean performance of our method does not depend, at the first order, on the distribution of the signal of interest $s(t)$.

In order to validate expression (10), we plot in figures 1, 2 and 3 the theoretical and experimental values of $1 - \rho/\rho_{max}$. Solid lines mean theoretical performance, '+','o' and 'x' mean experimental performance obtained by averaging over 50 independent trials. Figure 1 depicts $1 - \rho/\rho_{max}$ as a function of $N_1 = N_2$ for different values of λ. Figure 2 illustrates the behaviour of the method when $N_1 \neq N_2$: in this simulation, $1-\rho/\rho_{max}$ is plotted as a function of N_1 while N_2 is constant ($N_2 = 1000$). Figure 3 shows that the performance of the method is not modified when the source of interest is not Gaussian. In this figure, $N_1 = N_2$ and $s(t)$ is a constant modulus signal whose phase is uniformly distributed between 0 and 2π. These three figures allow not only to check the correctness of (10), but also to evaluate the area of validity of the asymptotical development.

4.3. Comparison with the optimal beamformer

We call *optimal beamformer* the spatial filter defined by $\hat{w} \stackrel{\text{def}}{=} (\widehat{R_1})^{-1} a$ which is commonly implemented when the source of interest steering vector a is known a priori. The calculations developed in the appendix, adapted to the case of $\hat{w}$, lead to the formula:

$$1 - \frac{\rho}{\rho_{max}} = \mu_1 \frac{M-1}{N_1} + o(N_1^{-1}) \tag{11}$$

which has to be compared with (10).

Notice that $\lambda \geq 1$ implies:

$$\left(\frac{1}{N_1} + \frac{1}{\lambda N_2}\right)\left(\frac{\lambda}{\lambda - 1}\right)^2 \geq \frac{1}{N_1} \tag{12}$$

It means that $\hat{w}$ always performs better than $\hat{v}$. This point is easy to understand intuitively: the a priori knowledge of a provides more information than the knowledge of an estimate of R_2.

But if σ_2^2 goes to infinity, then the knowledge of $\widehat{R_2}$ becomes "equivalent" to the knowledge of the vector a (up to a scalar constant). As a matter of fact, we also have:

$$\lim_{\sigma_2 \to +\infty} \left(\frac{1}{N_1} + \frac{1}{\lambda N_2}\right)\left(\frac{\lambda}{\lambda - 1}\right)^2 = \frac{1}{N_1} \tag{13}$$

If σ_2 and λ increase to infinity, then our filter provides asymptotically the same SINR than the optimal beamformer.

5. CONCLUSION

A blind source separation technique is proposed which improves existing methods by reducing the assumptions and increasing the performance. The computational cost is equivalent: one generalized eigendecomposition compared to two eigendecompositions. Additionally, the finite sample effect is calculated theoretically in a closed form expression and validated by simulations.

6. APPENDIX

We assume that the estimated filter $\hat{v}$ is equal to the optimal linear spatial filter v plus an error δv, i.e.:

$$\hat{v} \stackrel{\text{def}}{=} v + \delta v \tag{14}$$

By applying classical results of perturbation theory of the generalized eigendecomposition [3], or directly, by considering the equivalent definition (6) of $\hat{v}$, we obtain, at the first order:

$$\delta v = \mu_1 \frac{\lambda}{\lambda - 1} \Pi \left(\frac{\delta R_1}{\mu_1} - \frac{\delta R_2}{\mu_2} \right) v \tag{15}$$

with:

$$\Pi \stackrel{\text{def}}{=} R_n^{-1} - \frac{R_n^{-1} a a^H R_n^{-1}}{a^H R_n^{-1} a} \tag{16}$$

and, obviously: $\delta R_1 \stackrel{\text{def}}{=} \widehat{R}_1 - R_1$ and $\delta R_2 \stackrel{\text{def}}{=} \widehat{R}_2 - R_2$.

Then, the expectation of $\delta v \delta v^H$ is computed with the known expression of the covariance of the entries of the covariance matrices $\widehat{R}_1$ and $\widehat{R}_2$. This yields:

$$E(\delta v \delta v^H) = \mu_1 \, a^H R_n^{-1} a \left(\frac{1}{N_1} + \frac{1}{\lambda N_2} \right) \left(\frac{\lambda}{\lambda - 1} \right)^2 \Pi \tag{17}$$

Finally, by substituting this expression into (8) and with (9), we obtain (10).

7. REFERENCES

[1] Barry D. Van Veen and Kevin M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering", *IEEE ASSP Magazine*, pp. 4-24, April 1988.

[2] Mats Viberg, "Sensors Array Processing Using Gated Signals", IEEE ASSP, Vol. 37, No. 3, March 1989.

[3] Golub G. H. and C. F. Van Loan, "Matrix Computations", The Johns Hopkins University Press, 1989.

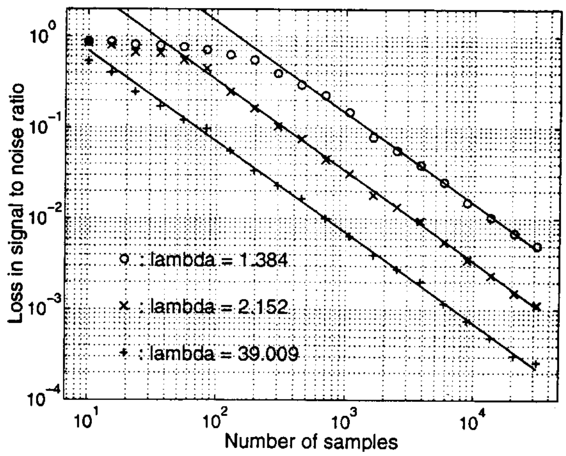

Figure 1: Theoretical vs. experimental performance.

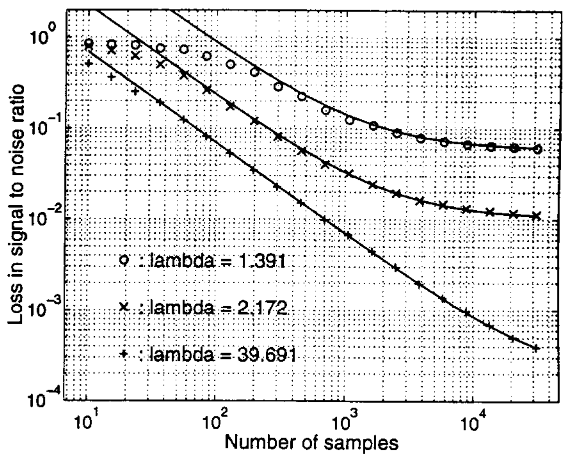

Figure 2: Theoretical vs. experimental performance.

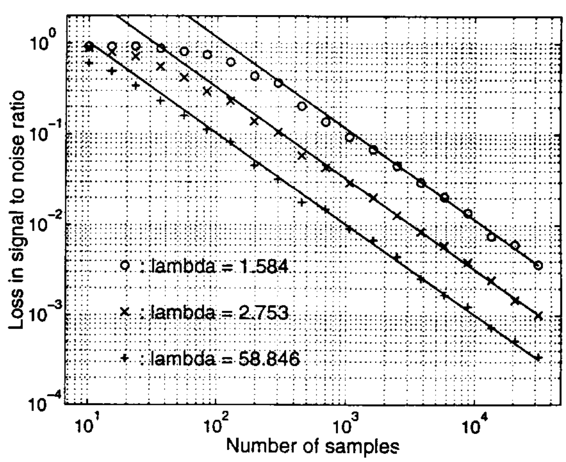

Figure 3: Theoretical vs. experimental performance.