Analytical Blind Discrete Source Separation

Olivier GRELLIER and Pierre COMON

Laboratoire I3S, Les Algorithmes/Euclide-B, 2000 route des Lucioles,
BP 121, F-06903 Sophia-Antipolis Cedex
grellier@i3sunice.fr comon@i3s.unice.fr

ABSTRACT

In this paper, a new blind source separation algorithm is described. Its main feature is to be **analytical**, in other words it doesn't suffer from local minima. The proposed method uses the discrete character of digital sources, which yields a polynomial system. The estimation of the sources is shown to be equivalent to the computation of rank-one tensors that are found by means of the old method by Macaulay for the computation of resultant. Finally, computer simulations are presented and the performances are compared to the analytical CM algorithm by van der Veen.

1 Introduction

Since the number of subscribers has impressively increased, blind source separation has become a crucial issue in the improvement on wireless communications. Indeed, blind source separation is one way of taking advantage of spatial diversity, without the help of learning sequences. Besides, thanks to blind techniques, the transmission rates could be increased in GSM or UMTS mobile systems. Static source separation is a first step towards the solution, which is indeed valid only in presence of flat fading.

Many algorithms have been proposed to solve blind source separation using various criteria. For example, one could use the independence of the sources [2] [1] or their constant modulus property [16] [12].

In this paper, we focus on the use of the discrete character of the sources. This approach has already been studied in [8] [4] [13], but our work is based on the seek for an **analytical solution**. In the BPSK case, Van der Veen proposed a partial analytical way to estimate the sources [15] (its solution used an iterative generalized Schur decomposition). Here, our solution is based on a new algorithm that analytically finds rank-one tensors.

2 Problem statement

2.1 Notation

Assume the following baseband reception model :

$$\mathbf{y} = H\,\mathbf{x} + \mathbf{w}, \tag{1}$$

where $\mathbf{x}$ is a random vector of size P, subsequently called the source vector, even if some components might be correlated to each other (*e.g.* multipaths), $\mathbf{w}$ is a random vector of size K standing for the noise, that will be assumed Gaussian and independent of $\mathbf{x}$. Lastly, $\mathbf{y}$ denotes the observation vector, A the mixing matrix, and K the number of sensors.

Finally, it is assumed throughout the paper that there are fewer sources than sensors, $K \geq P$.

2.2 Optimization criterion

When the source distribution is known and the noise is Gaussian, the optimal estimator in the MAP sense is given by :

$$(\hat{\mathbf{x}}, \hat{H})_{MAP} = \text{Arg}\min_{\mathbf{x}\in\mathcal{C},\hat{H}} ||\mathbf{y} - \hat{H}\hat{\mathbf{x}}||^2 \tag{2}$$

where $\mathcal{C}$ stands for the source constellation.

The computational complexity of this estimator is too heavy to be used, and we propose to minimize the following polynomial criterion [9],

$$\Phi(\mathbf{f}) = \frac{1}{N}\sum_{n=1}^{N}\prod_{x\in\mathcal{C}} |\mathbf{f}^{\mathsf{T}}\mathbf{y}(n) - x|^2$$

which has been shown to be asymptotically equivalent to the MMSE [6]. In this criterion, $\mathbf{f}$ stands for a spatial filter that extracts one discrete source and N is the number of observations.

2.3 Rank-one property

If the sources are PSK with D phase states (let's say PSK-D), the criterion Φ rewrites :

$$\Phi(\mathbf{f}) = \frac{1}{N}\sum_{n=1}^{N}\left|\left(\mathbf{f}^{\mathsf{T}}\mathbf{y}(n)\right)^D - 1\right|^2$$

and the spatial filter $\mathbf{f}$ must verify the following polynomial system :

$$\begin{cases} (\mathbf{f}^{\mathsf{T}}\mathbf{y}(1))^D &= 1 \\ &\vdots \\ (\mathbf{f}^{\mathsf{T}}\mathbf{y}(N))^D &= 1 \end{cases} \tag{3}$$

Using the symmetric vectorization operator, $\mathbf{vecs_D}\{.\}$, defined in [6] as a vector containing all the distinct cross-products of degree D appropriately weighted, system (3) becomes :

$$\begin{bmatrix} \mathbf{y}(1)^{\oslash D^{\mathsf{T}}} \\ \vdots \\ \mathbf{y}(N)^{\oslash D^{\mathsf{T}}} \end{bmatrix} \mathbf{f}^{\oslash D} \stackrel{\text{def}}{=} Y^{\oslash}\mathbf{f}^{\oslash D} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{4}$$

where $\mathbf{y}(n)^{\oslash D} = \mathbf{vecs_D}\{\mathbf{y}(n) \circ \cdots \circ \mathbf{y}(n)\}$ and $\mathbf{f}^{\oslash D} = \mathbf{vecs_D}\{\mathbf{f} \circ \cdots \circ \mathbf{f}\}$, and the operator $\circ$ stands for the tensor external product.

Since the spatial filter must verify systems (3) and (4), $\mathbf{f}$ is the mean squares solution of (3) under a structure constraint, which is such that the solution $\mathbf{f}^{\oslash D}$ must correspond to a rank-one tensor $\mathbf{F} = \mathbf{f} \circ \cdots \circ \mathbf{f}$.

3 Expression of tensor F

In the noiseless case, matrix $Y^{\oslash}$ is equal to

$$Y^{\oslash} \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{y}(1)^{\oslash D^{\mathsf{T}}} \\ \vdots \\ \mathbf{y}(N)^{\oslash D^{\mathsf{T}}} \end{bmatrix} = \begin{bmatrix} (H\mathbf{x}(1))^{\oslash D^{\mathsf{T}}} \\ \vdots \\ (H\mathbf{x}(N))^{\oslash D^{\mathsf{T}}} \end{bmatrix}$$

which rewrites

$$Y^{\oslash} = \begin{bmatrix} \mathbf{x}(1)^{\oslash D^{\mathsf{T}}} \\ \vdots \\ \mathbf{x}(N)^{\oslash D^{\mathsf{T}}} \end{bmatrix} H^{\oslash D^{\mathsf{T}}} \stackrel{\text{def}}{=} X^{\oslash D}\, H^{\oslash D^{\mathsf{T}}}$$

The rank of $Y^{\oslash}$ is then determined by the ranks of $H^{\oslash D}$ and $X^{\oslash D}$, assuming that the product introduces no loss of rank. It has been shown in [5], that this assumption is verified when $D = 2$. The proof for $D > 2$ may not be a problem but the calculus is more painful. In practice, *i.e.* during our simulations, this property has always been verified. Now, let's look at the rank of matrices $H^{\oslash D}$ and $X^{\oslash D}$.

$H^{\oslash D}$ can be expressed as $M\,(H \otimes H \otimes H \otimes H)\,N$, where $H \otimes H \otimes H \otimes H$ is full rank if H is full rank, and M and N are full rank matrices that transform $H \otimes H \otimes H \otimes H$ in a $K(K+1)(K+2)(K+3)/24$ times $P(P+1)(P+2)(P+3)/24$ matrix. Thus, if H is a square full rank matrix, then $H^{\oslash D}$ is also full rank. The investigations in other cases are more complex and have bot been carried out for the moment. However, if one wants to be sure that $H^{\oslash D}$ is full rank, it is sufficient to restrict the observation space to the signal subspace.

Now look at $X^{\oslash D}$. One of its line, say $\mathbf{x}(n)^{\oslash D^{\mathsf{T}}}$, contains all distinct Dth-degree monomials in $x_1(n)$, $x_2(n), \cdots, x_P(n)$. In each line, these monomials are listed in the same fixed order. Thus, if $\mathbf{x}_p$ is PSK-D, the column of $X^{\oslash D}$ corresponding to the monomial $x_p(n)^D$ has its entries equal to 1. If another source is PSK-D, another column of $X^{\oslash D}$ is made of ones and $X^{\oslash D}$ becomes singular. Therefore, if P sources are PSK-D, $X^{\oslash D}$ has a rank deficiency equal to $P-1$,

Since the product $X^{\oslash D}\, H^{\oslash D^{\mathsf{T}}}$ introduces no loss of rank, the $Y^{\oslash}$ kernel dimension is equal to $P-1$ if P PSK-D sources are present. Thus, the solution of (4) can be written as :

$$\mathbf{f}^{\oslash D} = \mathbf{f}^{\oslash}_{min} + \sum_{p=1}^{P-1} \lambda_p \mathbf{u}_p^{\oslash} \tag{5}$$

where $\mathbf{f}^{\oslash}_{min}$ is the pseudo-inverse solution.

Applying the inverse operator of $\mathbf{vecs_D}\{.\}$ to this equation we obtain :

$$\mathbf{F} = \mathbf{F}^{min} + \sum_{p=1}^{P-1} \lambda_p \mathbf{U}^p \tag{6}$$

where $\mathbf{F}$, $\mathbf{F}^{min}$ and $\mathbf{U}^p$ are respectively the D^{th}-order tensors associated with $\mathbf{f}$, $\mathbf{f}_{min}$ and $\mathbf{u}_p$.

4 Estimation of the sources

The key point of the proposed method is that $\mathbf{f}$ extracts a PSK-D signal if the D^{th}-order tensor $\mathbf{F}$ is rank-one. Thus, we must find the coefficients λ_p in (5) such that $\mathbf{F}$ is rank-one. In [16] van der Veen proposed an iterative algorithm that solves this problem for matrices. Here, we propose an **analytical** solution based on the Macaulay's method for the computation of resultants [10] [11], already used by the authors in the identification case [3].

4.1 Method for matrices ($D = 2$)

When $D = 2$, we are looking for a rank-one matrix F that is a linear combination of matrices F^{min} and U^p. In such a matrix, all 2×2 determinants vanish, *i.e.* the entries of $F = [f_{ij}]$ must verify :

$$f_{ij}\, f_{k\ell} = f_{kj} f_{i\ell} \quad \text{with} \quad k > i \text{ and } j < \ell \tag{7}$$

Replacing the expression (6) in the above equation leads to :

$$\left(f^{min}_{ij} + \sum_{p=1}^{P-1} \lambda_p\, u^p_{ij}\right)\left(f^{min}_{k\ell} + \sum_{p=1}^{P-1} \lambda_p\, u^p_{k\ell}\right) = \left(f^{min}_{kj} + \sum_{p=1}^{P-1} \lambda_p\, u^p_{kj}\right)\left(f^{min}_{i\ell} + \sum_{p=1}^{P-1} \lambda_p\, u^p_{i\ell}\right)$$

which provides a system of polynomial equations of second degree that can be solved using the Macaulay's method described in [3].

In our case, Bézout's theorem [7](p.227) shows that the polynomial system has at most 2^{P-1} solutions. Then, the analytical method used in [3] returns 2^{P-1} possible solutions, among which only P give rank-one matrices F as we show now.

If P sources are present, there exist P spatial filters $\mathbf{f}_p$ that satisfy equations (5) and (6). Hence, there are at least P rank-one solutions among the 2^{P-1} solutions. Suppose there is one more rank-one solution. Since it also satisfies equation (6), the corresponding spatial filter extracts a PSK-D source, that is already extracted (see the unicity results in [14]). Hence, there may exist 2 spatial filters $\mathbf{f}_1$ and $\mathbf{f}_2$ that extract the same source :

$$\begin{aligned} \mathbf{f}_1^{\mathsf{T}} Y = \mathbf{f}_2^{\mathsf{T}} Y \quad & \Leftrightarrow \quad (\mathbf{f}_1 - \mathbf{f}_2)^{\mathsf{T}} Y = 0 \\ & \Leftrightarrow \quad dim\left(Ker\left(Y\right)\right) \neq 0 \end{aligned}$$

which is impossible if each sensor receives a distinct mixture of the sources.

Therefore, the steps of our algorithm are :

1. Restrict the observation to the signal subspace,
2. Compute the pseudo-inverse of (4) and the kernel of $Y^{\oslash}$
3. Compute all the 2×2 determinants,
4. Solve the polynomial system
5. Finally, keep the P solutions nearest to a rank-one matrix.

4.2 Extension to tensors ($D > 2$)

When $D > 2$, we must find D^{th}-order tensors that are rank-one. The matrix slices of these tensors must all be rank-one. Then, the algorithm described for $D = 2$ also works in this case and gives P source estimates. The algorithm becomes :

1. Restrict the observation to the signal subspace,
2. Compute the pseudo-inverse of (4) and the kernel of $Y^{\oslash}$
3. For each matrix slice, compute all the 2×2 determinants and solve each polynomial system separatly
4. Finally, keep the P solutions nearest to a rank-one tensor among all the solutions found.

5 Computer results

The simulations have been carried out using a uniformly spaced linear array with $K = 4$ sensors. The element spacing is $\lambda/2$, where λ is the wavelength of the propagating waveforms. $P = 3$ uncorrelated sources impinge on the array, and the directions of arrival are $\theta_1 = -20°$, $\theta_2 = 3°$ and $\theta_3 = 15°$. The proposed algorithm has been tested with modulation QPSK and $N = 300$ observations.

Figures 1, 2 and 3 show the average Bit Error Rates obtained over 500 trials and for various SNRs. The performances of our algorithm are compared to the LS spatial filter (dashed dot), computed assuming that the sources are completly known, and compared to the performances of the ACMA, Analytical CM Algorithm, (dashed), proposed by Van der Veen in [16].

These simulations show that the performances of our analytical algorithm are close to those of ACMA, compared to the LS performances. However, the use of the discrete character of the sources allow a decrease in the BER when the SNR is sufficiently low. Indeed, at 10 dB, our performances are under the performances of ACMA. This is due to the fact that our approach is based on an approximation of the MMSE which does not hold when the SNR is too low.

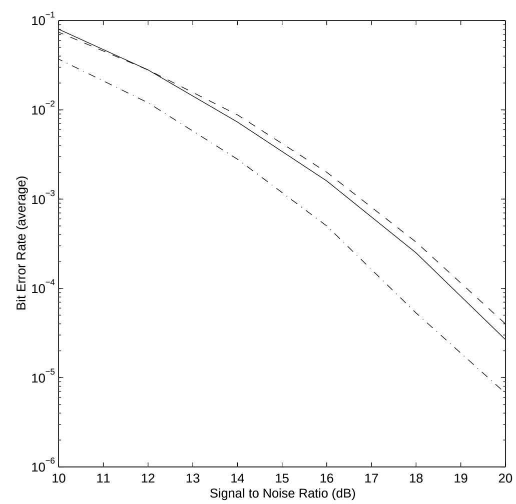

Figure 1: Computer results for the first QPSK source signal.

6 Conclusion

In this paper, we proposed a new approach to discrete source separation. We extended the algorithm of van der Veen in [15] to PSK sources by introducing a symmetric tensor vectorization operator $\mathbf{vecs_D}\{.\}$. The estimation of the sources is performed by looking for rank-one tensors that are a linear combination of given tensors. This is done entirely **analytically**.

The performances of the new algorithm show that the use of the discrete character of the sources allow a decrease in BER when the SNR is not too low.

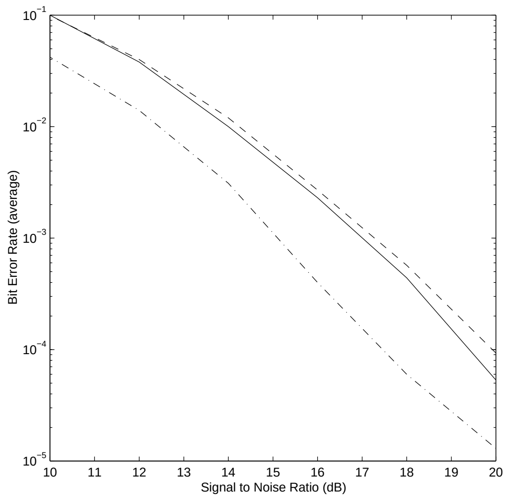

Figure 2: Computer results for the second QPSK source signal.

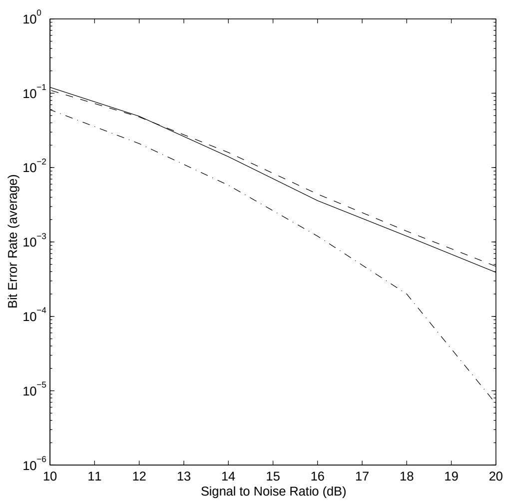

Figure 3: Computer results for the third QPSK source signal.

References

[1] J. F. CARDOSO, A. SOULOUMIAC, "Blind beamforming for non-Gaussian signals", *IEE Proceedings - Part F*, vol. 140, no. 6, pp. 362–370, Dec. 1993, Special issue on Applications of High-Order Statistics.

[2] P. COMON, "Independent Component Analysis, a new concept ?", *Signal Processing, Elsevier*, vol. 36, no. 3, pp. 287–314, Apr. 1994, Special issue on Higher-Order Statistics.

[3] P. COMON, O. GRELLIER, B. MOURRAIN, "Closed-form blind channel identification with MSK inputs", in *Asilomar Conference*, Pacific Grove, California, November 1-4 1998, pp. 1569–1573, Invited session.

[4] F. GAMBOA, "Separation of sources having unknown discrete supports", in *IEEE-ATHOS Workshop on Higher-Order Statistics*, Begur, Spain, 12–14 June 1995, pp. 56–60.

[5] O. GRELLIER, *Déconvolution et séparation de sources discrètes*, Doctorat, Université de Nice Sophia-Antipolis, janvier 2000.

[6] O. GRELLIER, P. COMON, "Closed-form equalization", in *SPAWC*, Annapolis, MD, May 9-12 1999, pp. 219–222.

[7] J. HARRIS, *Algebraic Geometry, a first course*, vol. 133 of *Graduate Texts in Math.*, Springer, 1992.

[8] A. KANNAN, V. U. REDDY, "Maximum likelihood estimation of constellation vectors for blind separation of co-channel BPSK signals and its performance analysis", *IEEE Trans. Sig. Proc.*, pp. 1736–1741, 1997.

[9] T.-H LI, "Blind identification and deconvolution of linear systems driven by binary random sequences", *Trans. on Inf. Theory*, vol. 38, no. 1, pp. 26–38, Jan 1992.

[10] F. S. MACAULAY, "Some formulae in elimination", *Proc. London Math. Soc.*, vol. 1, no. 33, pp. 3–27, 1902.

[11] B. MOURRAIN, "Computing isolated polynomial roots by matrix methods", *J. Symbolic Computation, Special Issue on Symbolic-Numeric Algebra for Polynomials*, Dec. 1998.

[12] A. J. PAULRAJ, C. PAPADIAS, "CM algorithm for multiuser signal separation in presence of delay spread using antenna arrays", *IEEE Sig. Proc. Letters*, vol. 4, no. 6, pp. 178–181, June 1997.

[13] K. MBAREK T.-H LI, "A blind equalizer for nonstationnary discrete valued signals", *Trans. on Sig. Proc.*, vol. 45, no. 1, pp. 247–254, Jan 1997.

[14] S. TALWAR, M. VIBERG, A. PAULRAJ, "Blind estimation of multiple co-channel digital signals arriving at an antenna array: Part I, algorithms", *IEEE Trans. Sig. Proc.*, pp. 1184–1197, May 1996.

[15] A. J. van der VEEN, "Analytical method for blind binary signal separation", *IEEE Trans. Sig. Proc.*, vol. 45, no. 4, pp. 1078–1082, Apr. 1997.

[16] A. J. van der VEEN, A. PAULRAJ, "An analytical constant modulus algorithm", *IEEE Trans. Sig. Proc.*, vol. 44, no. 5, pp. 1136–1155, May 1996.