# FEASIBILITY OF SOURCE SEPARATION IN FREQUENCY DOMAIN

*Ch. Servière*

CEPHAG-ENSIEG BP 46
38402 Saint-Martin d'Hères, FRANCE

## ABSTRACT

We focus on the feasibility of the source separation in the frequency domain. First, it is linked with the convergence speed towards gaussianity of signals after L-point discrete Fourier Transform. We test here a distance to gaussianity thanks to the spectral kurtosis. We analyse the influence of L, of the duration of the source tricorrelations and of a non linear filtering. We mainly develop the case of QARMA processes. The second point consists in the reconstruction of the spectra of the estimated sources from the signals identified at each frequency bin. Indeed, the source associated to the ith identified signal is not necessarily the same from one frequency bin to another. The algorithm efficiency is then illustrated on QARMA processes, including the procedures of separation and reconstruction.

## 1. INTRODUCTION

The blind source separation issue has recently but largely been investigated as it arises in many fields (noise reduction, radar and sonar processing, speech enhancement, separation of rotating machine noises, localization in array processing, ...). Whatever its application, it consists in recovering the signals emitted by p sources $\underline{s}(t)$ from M observed linear mixtures $\underline{r}(t)$ of these sources. In a general context of source separation, the only three assumptions are the non-gaussianity of the source signals, their mutual independence and the linearity and stationarity of the propagation.
Since ten years, many solutions have been proposed which test different measurements of the statistical independence. They are based on the use of fourth-order moments or cumulants, nonlinear functions, contrast functions or the information maximization principle [1] [2] [3] [4] [5] [6]. The linear filters (which characterize the propagation from the sources to the sensors) can be estimated using adaptive or nonadaptive algorithms minimizing or looking for zeros of different independence criteria. Some procedures are also based on the maximum likelihood principle [4] [6].
Most of these works reconstruct the source signals from instantaneous mixtures. In a general case of convolutive mixtures, the problem has been treated with the help of adaptive algorithms in the time domain [3]. Certain symmetrical fourth-order cumulants of the estimated sources are proposed to be canceled. Yet, it is proved that it is a sufficient condition to separate the sources only under the hypothesis of independent, identically distributed (i.i.d) processes with the same sign of kurtosis. In many applications, the statistical properties of the source signals are far from the previous hypotheses which achieve the separation. Besides, several problems (spurious solutions, local minima, influence of the initialization and low convergence speed) prevent from using these methods.
The source separation problem may also be treated in the frequency domain thanks to the use of multispectra [13]. In this approach, the general problem of convolutive mixtures comes to a problem of instantaneous mixtures of narrow-band sources in each frequency band. The mixtures in each band are then separated using one of the previous methods for complex mixtures.
We propose in this paper a study of the source separation problem to convolutive mixtures of wide-band sources in the frequency domain. We mainly develop two points. The feasibility is linked with the convergence speed towards gaussianity of the signals after L-point discrete Fourier Transform. In §3, we test a distance to gaussianity thanks to the spectral kurtosis. We analyse the influence of L, of the duration of the source tricorrelations and of a non linear filtering. We give then a lower bound for the spectral kurtosis, available for the whole frequency band. After that, we mainly develop the case of QARMA processes in §3.4.
The second point consists in the reconstruction of the estimated sources spectra from the signals identified at each frequency bin, as the source associated to the ith identified signal is not necessarily the same from one frequency bin to another. We discuss this point in §4 and propose an original method to solve the problem. The algorithm efficiency is then illustrated on QARMA processes, including the procedures of separation and reconstruction.

## 2. MODELIZATION OF THE PROBLEM

In a general blind source separation problem, the observed M-dimensional data vector $\underline{r}(t)$ may be represented in frequency-domain by an instantaneous complex mixture for each frequency bin n, which leads to the following model:

(1) $\underline{Ri}(n) = \underline{\underline{A}}(n)\ \underline{Si}(n) + \underline{Vi}(n) \quad n=0, ..., L-1$

where $\underline{Ri}(n)$ is the L-point Discrete Fourier Transform (DFT) of the ith data block of the observation $\underline{r}(t)$. $\underline{Si}(n)$ represents the DFT of the ith data block of the p-dimensional data vector of the sources $\underline{s}(t)$. $\underline{\underline{A}}(n)$ is a matrix (M.p) which characterizes the linear propagation from sources to sensors and $\underline{Vi}(n)$ represents an additive M-dimensional gaussian noise. The problem consists first in identifying the matrix $\underline{\underline{A}}(n)$. After a singular value decomposition, the mixing matrix $\underline{\underline{A}}(n)$ is expressed as the product of three matrices.

(2) $\underline{\underline{A}}(n) = \underline{\underline{F}}(n)\ \underline{\underline{D}}(n)\ \underline{\underline{\Pi}}(n)$

where $\underline{\underline{F}}(n)$ and $\underline{\underline{\Pi}}(n)$ are two (M.M) and (p.p) unitary matrices. $\underline{\underline{D}}(n)$ is a (M.p) diagonal matrix. The two matrices

$\underline{\underline{F}}(n)$ and $\underline{\underline{D}}(n)$ are identified thanks to second-order statistic criteria. They respectively contain the eigenvectors and the eigenvalues of the spectral matrix of the observation $\underline{Ri}(n)$.

After projection of the observation vector $\underline{Ri}(n)$ in the signal subspace (which is spanned by the eigenvectors associated to the dominant eigenvalues) and normalization, the components of the p-dimensional vector, noted $\underline{Xi}(n)$, are uncorrelated and normalized. They are relied to the components of the normalized source vector, noted $\underline{s}'(t)$, by:

(3) $$\underline{Xi}(n) = \underline{\underline{\Pi}}(n)\ \underline{S'i}(n)$$

where $\underline{S'i}(n)$ is the DFT of $\underline{s}'(t)$.

$\underline{\underline{\Pi}}(n)$ may be expressed as a product of Givens rotations and estimated thanks to fourth-order criteria, by testing different measurements of the statistical independence, as presented in §1.

# 3 STUDY OF THE SPECTRAL KURTOSIS

## 3.1 Computation of a lower bound

This part of the paper is devoted to the convergence speed towards gaussianity of the signals after L-point discrete Fourier Transform. As it has been shown in §2, the source separation methods lay on the additional information provided by fourth-order statistics. This information only exists under the hypothesis of non gaussian sources. Whatever the chosen method, the variance of the estimator of matrix $\underline{\underline{\Pi}}(n)$ is inversely proportional to the kurtosis of the sources [8]. In a similar way, in frequency domain, we study the distance of the DFT to gaussianity thanks to the spectral kurtosis which is defined as a section of the general trispectrum of the normalized sources. Let K(Si(n)) be the kurtosis of Si(n), defined by :

(4) $$K(Si(n)) = \frac{cum(Si(n), Si(n)^*, Si(n), Si(n)^*)}{cum(Si(n), Si(n)^*)}$$

where cum represents the cumulants of second and fourth orders and * the complex conjugate.

In [9], a generalized central limit theorem is presented under a sufficient condition of convergence relative to the duration of the multicorrelation of the sources. Suppose s(t) is a strictly stationary process, all of whose moments exist, we define its L-point DFT by :

(5) $$S(n) = \sum_{m=0}^{L-1} s(m) \exp\left(\frac{-2\pi jnm}{L}\right)$$

$C_S k(u_1, \ldots, u_{k-1})$ represents the source multicorrelation of order k. If the span of dependence of s(t) is small enough that :

(6) $$\sum_{u_1, \ldots, u_{k-1}}^{+\infty} \left|C_S k(u_1, \ldots, u_{k-1})\right| < +\infty, k = 2, \ldots, +\infty$$

then the variables S(n) are asymptotically independent and have asymptotically a complex normal distribution [9], with zero mean and variance (L fs(n)), when L tends towards infinity. fs(n) represents the power spectrum of s(t).

For finite values of L, the convergence speed is linked with the duration of the source multicorrelations. We test here a distance to gaussianity thanks to the spectral kurtosis. We compute then a lower bound for it, which is available in the whole frequency-band. From (4), we remark that the spectral kurtosis in unchanged by linear filtering. We can then suppose that cum(Si(n),Si(n)*) is constant. Now suppose that the duration of the tricorrelation of s(t) is bounded and equal to T. After computations, the numerator of the K(Si(n)) is then equal to : (7)

$$N(n) = \Big( \sum_{k=-T+1}^{T-1} \sum_{l=-T+1}^{T-1} \sum_{m=-T+1}^{T-1} [\ L - |\max(k,l,m,0)| - |\min(k,l,m,0)|].\ C_S 4(k,l,m) \exp\left(\frac{-2\pi jn(k-l+m)}{L}\right)\Big)$$

Denote $\varepsilon$ : (8)

$$\varepsilon = \Big| \sum_{i=0}^{L-1} \sum_{\substack{k,l,m=-i \\ (k-l+m=0)}}^{L-1-i} C_S 4(k,l,m) - \sum_{i=0}^{L-1} \sum_{\substack{k,l,m=-i \\ (k-l+m\neq 0)}}^{L-1-i} C_S 4(k,l,m) \Big|$$

with |k|, |l|, |m| <T

From (8), it can be proved that $(|N(n)|^2 - \varepsilon^2)$ is always positive. Consequently, |K(S(n))| admits the lower bound $(\varepsilon/(L^2 C_S 2(0)^2)$, independent of the frequency bin n. The lower bound of the spectral kurtosis depends both on T and on the shape of the source tricorrelation. It also shows the distinct influence of the tricorrelation terms for (k-l+m=0) and (k-l+m≠0).

This lower bound is very useful to know the feasibility of the source separation on the whole frequency band. In several application, like active sonar environment, the sources (or at least one of the sources) are controlled. This lower bound makes it possible to test the spectral kurtosis on the whole frequency band. We show on some examples §3.4 that it is able to construct signals such that the spectral kurtosis value is significant. We analyse in particular the influence of nonlinear filtering on the spectral kurtosis value in §3.2, including for example QARMA processes (§3.4).

## 3.2 Influence of nonlinear filtering on the spectral kurtosis

Suppose $Y_i(f)$, the Fourier transform of the data block [y(i), ..., y(i+L-1)] of continuous frequency f, where y(i) is a nonlinear filtering of s(i). Denote K(Yi(f)), the spectral kurtosis of Yi(f). We show that K(Yi(f)) cannot be strictly equal to K(Si(f)) if y(t) is strictly a nonlinear filtering of s(t).

A general model for $|Yi(f)|^2$ is given by :

(9) $$|Yi(f)|^2 = |H(f)|^2\ |Si(f)|^2 + Zi(f)$$

where Zi(f) is a real stationary random sequence of time index i, correlated with Xi(f). $|H(f)|^2\ |Si(f)|^2$ represents the linear part of the filtering between Si(f) and Yi(f). We assume that $E\left\{Zi(f)|Si(f)|^2\right\}$ is non zero. K(Yi(f)) is equal to K(Si(f)) if and only if it exists positive roots of equation (10) :

$$|H(f)|^4 + 2|H(f)|^2 \frac{E\left\{Zi(f)|Si(f)|^2\right\}}{E\left\{|Si(f)|^4\right\}} +$$

$$\frac{E\{|Zi(f)|^2\}}{E\{|Si(f)|^4\}} - \frac{E\{|Yi(f)|^2\}^2}{E\{|Si(f)|^2\}^2} = 0$$

It can be proved from (10) that the roots are strictly positive. Consequently, if y(t) is strictly a nonlinear filtering of s(t) (H(f)=0), K(Yi(n)) increases or decreases versus K(Si(f)) in the whole frequency band. It increases or decreases according to the type of nonlinearity. It depends if the kurtosis of s(t) is inferior (or superior) to the kurtosis of y(t).

The difference between K(Si(f)) and K(Yi(f)) is necessarily linked to the shapes of the autocorrelations and the tricorrelations of y(t) and s(t). It can be approximated by the lower bounds (8).

### 3.3. Application to QARMA processes

QARMA processes are obtained by squaring ARMA processes. They are generated as described in (11):

(11) $z(m)=y(m)^2-E\{y(m)^2\}$

where y(m) is an ARMA process issued from x(m). z(m) is a strictly nonlinear filtering of x(m). As explained before, K(Z(n)) increases versus K(X(n)) if and only if K(z) is superior to K(x). Now suppose that x(m) is a gaussian process. The above condition is necessarily verified.

After some computations, we obtain that :

(12) $C_Z4(k,l,m)=16[\Gamma(k)\Gamma(l)\Gamma(k-m)\Gamma(l-m)+\Gamma(k)\Gamma(m)\Gamma(k-l)\Gamma(l-m)+\Gamma(l)\Gamma(m)\Gamma(k-l)\Gamma(k-m)]$

(13) $C_Z2(k)=2\Gamma(k)^2$

where $\Gamma(k)=E\{y(t)y(t-k)\}$. Let y(m) be an AR1 gaussian process. Denote its autocorrelation $\Gamma(k)$ :

(14) $\Gamma(k)=\alpha^{|k|}\ \Gamma(0)$ with $0 < \alpha < 1$

From (12) (13) (14), we can prove after some computations that K(Z(n)) increases with the coefficient $\alpha$. We show in figure 1 the spectral kurtosis for two values of $\alpha$ (0.8;0.9) (L=64).

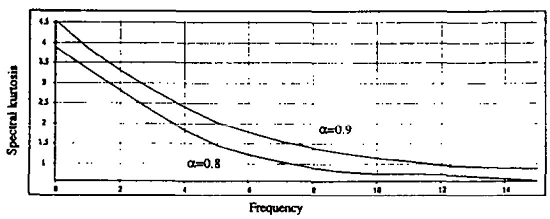


Fig. 1 : Spectral Kurtosis for ($\alpha$ =0.8 and $\alpha$ =0.9)

We verify that K(Z(n)) is correctly increasing with $\alpha$. The values of K(Z(n)) make it possible to apply source separation methods on such signals. Experimental results on this type of signals are shown in §5, including the procedures of separation and reconstruction .

# 4. RECONSTRUCTION OF THE SOURCE SPECTRA

The crucial point consists in the reconstruction of the time sources sk(t) for (k=1, ..., p). After identification of the matrix $\underline{\underline{\Pi}}(n)$ with fourth-order criteria, p independent components of $\underline{S}i(n)$ are extracted in each frequency bin n. More precisely, the identified matrix, noted $\hat{\underline{\underline{\Pi}}}(n)$, is relied to $\underline{\underline{\Pi}}(n)$ by :

(15) $\hat{\underline{\underline{\Pi}}}(n)=\underline{\underline{\Pi}}(n)\underline{\underline{P}}(n)\underline{\underline{\Delta}}(n)$

where $\underline{\underline{P}}(n)$ is a (p.p) permutation matrix and $\underline{\underline{\Delta}}(n)$ is a diagonal one. Due to the existence of this permutation matrix at each bin n, and since the methods independently treat each frequency bin, the kth identified component of $\underline{S}i(n)$ is not necessarily associated to the same time source sk(t), from one frequency bin to another.

In order to re-establish the continuity of the source spectra, several ideas have been investigated. The first one consists in examining the statistic relationship between the different estimated sources from one frequency bin to another. However, the correlation and the tricorrelation between two adjacent frequency bins of the same source depend on the shape of their spectral density and trispectrum [10]. It is sometimes impossible to define an absolute threshold from which the frequency components of two estimated signals would be attributed to the same temporal source or to two different sources. This remark shows the interest of going back on a temporal expression of the sources.

The proposed method aims first at recovering the statistic relationship between the estimated sources at one frequency-bin n, $\underline{S}i(n)$, and the temporal sources $\underline{s}(t)$.

Recall the expression of the DFT of the kth component of $\underline{S}i(n)$, Si(n,k), on the ith data block :

(16) $$Si(n,k)=\sum_{l=0}^{L-1} sk(i+l)\exp(-2\pi jn\frac{l}{L})$$

Recall the expression of the DFT of the kth component of $\underline{S}i(n)$, on the (i+1)th data block, delayed for one sample :

(17) $$Si+1(n,k)=\sum_{l=0}^{L-1} sk(i+1+l)\exp(-2\pi jn\frac{l}{L})$$

We easily deduce from the two expressions (16) (17) a relation between an estimated source in frequency domain Si(n,k), at data blocks i and (i+1), and the associated temporal source sk(t): (18)

$$Si+1(n,k)\exp(-2\pi j\frac{n}{L})-Si(n,k)=sk(i+L)-sk(i)$$

which directly ensues from the expression of the DFT. Consequently, it exists a specific MA filtering of $\underline{S}i(n)$ for each frequency bin n which is relied to a function of the time sources ($\underline{s}(t)$ - $\underline{s}(t-L)$).

Suppose now that the components of index i0, Si(n,io), and j0, Si(n+1,jo), are associated to the same time source s1(t). The previous MA filtering of Si(n,io), H(Si(n,io)), and of Si(n+1,jo), H(Si(n+1,jo)), will be equal to the same quantity (s1(i+L-1)-s1(i-1)), whatever the treated frequency bin n. Consequently, their coherence will be non zero and equal to one. On the contrary, if the two components are not associated to the same time source, H(Si(n,io)) H(Si(n+1,jo)), will be uncorrelated. Their coherence will then be zero. As a conclusion, a criterion based on the second order moments may be used to associate the L frequency components of s1(t):

(19) $E\{H(Si(n,k)).H(Si(n+1,j))^*\}\neq 0$ for $(k,j)=(1,1)$

$=2\ [\Gamma s1(0)-\Gamma s1(L)]$

where $\Gamma s1(k)$ represents the autocorrelation of the source s1(t).

$$(20)\ E\{H(Si(n,k)).H(Si(n+1,j))^*\} = 0 \text{ for } (k,j) \neq (1,1)$$

This criterion is ambiguous in the only case of periodic sources of period L, which is quite a particular case.

Different algorithms of reconstruction are possible, depending on the class of signals. If the spectra of the sources are continuous, the reconstruction may be realized between each frequency bin and its previous bins. The contrast of the coherence between H(Si(n,io)), H(Si(n,jo)), and H(Si(n+1,k0)) allows to associate the k0-th component (at bin n+1) with the component i0 or j0 at bin n. This procedure is illustrated in §5 fig.3 on experimental data.

## 5. EXPERIMENTAL RESULTS

The simulation here after illustrates the results of a complete implementation of the algorithm in the frequency-domain, in the case of convolutive mixtures of two sources. The two sources are QARMA processes, mixed in a noisy context. We present in figure 2 the coherences between the two observations and the two sources. It proves that the sources are actually mixed in the observations. After separation in frequency domain, two signals are identified in each frequency bin. They have been extracted, using a likelihood principle [4]. We present in figure 3 the coherences between the two estimated sources after filtering by the specific filters proposed in §4 H(Si(n,1)), H(Si(n,2)), and one source estimated at bin n+1, H(Si(n+1,1)). We remark that, for each frequency bin, one estimated source is correctly correlated with S1(n) since the second one is uncorrelated with it. The contrast of the coherence between H(Si(n,1)), H(Si(n,2)), and H(Si(n+1,k0)) allows to associate the k0-th component (at bin n+1) with the component 1 or 2 at bin n. This shows a very good quality of the separation procedure in all the frequency band. We present in figure 4 the coherences between the two right sources and the estimated ones after separation and reconstruction using the proposed technique in §4. The two curves close to 1 and the two curves close to 0 reveal a good quality of separation for each frequency bin and a good quality of reconstruction of the sources.

## 6. CONCLUSION

We propose in this paper a generalization of the source separation problem to convolutive mixtures of wide-band sources in the frequency domain. We develop two specific points. In the first point, we study the feasibility of the separation in frequency domain with regard to the hypothesis of non gaussian sources. It is linked with the convergence speed towards gaussianity of signals after L-point discrete Fourier Transform. We test here a distance to gaussianity thanks to the spectral kurtosis. We analyse the influence of L, of the duration of the source tricorrelations and of a non linear filtering. We mainly develop the case of QARMA processes. The second point consists in the reconstruction of the estimated sources spectra from the signals identified at each frequency bin, since the sources associated to the ith identified signals are not necessarily the same from one frequency bin to another. The algorithm efficiency is then illustrated on QARMA processes, including the procedures of separation and reconstruction.

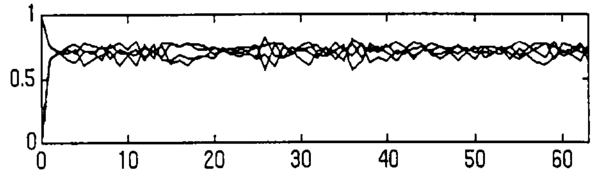


Fig2. Coherences between the observations and the right sources.

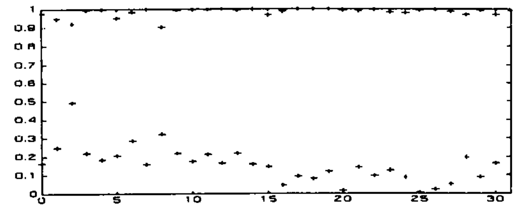

Fig3. Coherences between H(Si(n,1)), H(Si(n,2)) and one estimated source in the reference bin 16, H(Si(16,1)).

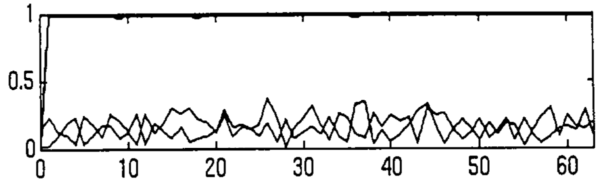


Fig4. Coherences between the right sources and the estimated ones


**REFERENCES**

[1] P.COMON, "Independent component analysis, a new concept?", *Signal Processing*, vol 36, n°3, April 1994, pp 287-314

[2] J. F. CARDOSO, A.BELOUCHRANI, B.LAHELD, "A new composite criterion for adaptive and iterative blind source separator",*Proc ICASSP*, vol 4, pp 273-276, 1994.

[3] H.L. NGUYEN THI, Ch. JUTTEN, "Blind sources separation for convolutive mixtures", *Signal Processing*, vol 45, n° 2, August 1995, pp 209-229

[4] M. GAETA, J. L. LACOUME, "Source separation versus hypothesis",*Workshop on Higher Order Statistics*, Chamrousse, France, pp 271-274, 1991.

[5] A.BELL, T.SEJNOWSKI,"An information-maximization approach to blind separation and blind deconvolution", *Neural Computation*, vol 7, pp 1004-1034, 1995

[6] D.T. PHAM, P.GARAT; Ch. JUTTEN, "Separation of a mixture of independent sources through a maximum likelihood approach", in *Proc. EUSIPCO-92*, Brussels, August 1992, pp 771-77

[7] D.YELLIN, E.WEINSTEIN, "Multi-Channel Signal Separation based on Cross-Bispectra", *Proc. Int. Workshop on Higher Order Statistics*, Lake Tahoe, pp 270-274, 1993

[8] P. CHEVALIER, "On the performances of higher order blind sources separation methods", *Proc. Int. Workshop on Higher Order Statistics*, Girona, pp.30-35, 1995

[9] D.BRILLINGER,"Time Series, Data analysis and Theory", Holden-Day, 1981, p94

[10] V.CAPDEVIELLE, "Séparation de sources large bande à l'aide de moments d'ordre supérieur", *Ph.D. Thesis*, INPGrenoble, 1995