Séparation de Mélanges Convolutifs par Diagonalisation Conjointe

Pierre Comon[1], Éric Moreau[2], Ludwig Rota[1]

[1]Labo. I3S, CNRS/UNSA, BP 121, 06903 Sophia-Antipolis cedex

[2]Labo. SIS, ISITV, BP 56, 83162 La Valette du Var cedex

comon@i3s.unice.fr moreau@isitv.univ-tln.fr

Résumé – La séparation aveugle de mélanges convolutifs et l'égalisation aveugle de canaux à entrées et sorties multiples sont un seul et même sujet, que nous abordons ici. L'algorithme numérique, présenté en détail, s'appuie sur des résultats théoriques sur les contrastes publiés récemment par les auteurs. Cet algorithme consiste en la diagonalisation partielle approximative conjointe de plusieurs matrices contenant certaines valeurs des multi-corrélations cumulantes des observations.

Abstract – Blind Separation of convolutive mixtures and Blind Equalization of Multiple-Input Multiple-Output (MIMO) channels are two different ways of naming the same problem, which we address here. The numerical algorithm, subsequently presented in detail, is based on theoretical results on contrasts recently published by the authors. This algorithm consists of Partial Approximate Joint Diagonalization of several matrices, containing some values of output cumulant multi-correlations.

1 Introduction

Le problème de l'égalisation *aveugle* (càd sans connaître les entrées) de systèmes linéaires invariants dans le temps a fait l'objet de travaux approfondis depuis une douzaine d'années.

Parmi les travaux antérieurs, on peut citer [2] [11] pour l'égalisation SISO avec les Statistiques d'Ordre Élevé (SOE), ou [17] [9] avec le critère du module constant, ou encore [7] avec le critère de la puissance constante. Dans le cas de canaux multiples, on peut mentionner [12] [1] [14] [6] [13]. D'autres travaux ont été proposés pour des algorithmes en-ligne, parmi lesquels [15], certains procédant par déflation [16]. Le cas des mélanges statiques de sources a aussi fait couler récemment beaucoup d'encre. Sans citer les travaux les plus récents, on peut dire qu'il y a deux approches; la première [5] procède par diagonalisation d'un tenseur, et la seconde [3] par diagonalisation approchée de certaines tranches de ce même tenseur. Ces approches originales étaient de type "bloc" (hors-ligne). Ce modèle statique n'est utilisable sur le plan pratique en communications numériques que lorsque les retards d'arrivée des trajets dans un canal spéculaire sont inférieurs à la durée symbole, ce qui est finalement assez restrictif.

Notre contribution principale est de proposer un algorithme *analytique* pour l'égalisation MIMO (entrées et sorties multiples); en d'autres termes, nous séparons des mélanges convolutifs de plusieurs sources en présence de plusieurs capteurs. L'égalisation est mise en œuvre à l'aide d'un algorithme original *hors-ligne*, dont il est aisé de construire une version *en-ligne* temps réel si nécessaire. En outre, la solution maximise une fonction de *contraste*, et est donc *optimale* à ce titre.

Dans [8], nous avions développé les fondements théoriques. Ici, nous présentons l'algorithme numérique et les résultats de simulation, qui montrent son efficacité sur des blocs **très courts**.

2 Position du problème

Le système linéaire inversible considéré suit le modèle suivant:

$$\boldsymbol{x}(n) = \sum_{k=-\infty}^{\infty} \boldsymbol{C}(k)\boldsymbol{a}(n-k) \tag{1}$$

où $\boldsymbol{a}(n)$ est le vecteur source de dimension N, $\boldsymbol{x}(n)$ est le vecteur des observations, supposé de même dimension, et $\{\boldsymbol{C}\} \stackrel{\text{def}}{=} \{\boldsymbol{C}(n), n \in \mathbb{Z}\}$ est la suite des matrices $N \times N$ définissant la réponse impulsionnelle du canal (pour plus de clarté, les vecteurs sont notés en minuscules grasses, et les matrices en majuscules grasses). L'égalisation aveugle consiste à trouver le filtre $\{\boldsymbol{H}\} \stackrel{\text{def}}{=} \{\boldsymbol{H}(n), n \in \mathbb{Z}\}$ uniquement à partir des observations $\boldsymbol{x}(n)$.

Pour ce faire, on fait les hypothèses suivantes: (**H1**) Les sources $a_i(n)$, $i \in \{1, \ldots, N\}$, sont indépendantes, et chaque source est une séquence i.i.d. centrée réduite. (**H2**) $\boldsymbol{a}(n)$ est un processus stationnaire jusqu'à l'ordre considéré, $r \in \mathbb{N}^*$, *i.e.* $\forall i \in \{1, \ldots, N\}$, le cumulant d'ordre r $\mathsf{Cum}\,[a_i(n), \ldots, a_i(n)]$ est indépendant de n, et sera noté $\mathsf{C}_r[a_i]$. (**H3**) Au plus une des sources est de cumulant nul. (**H4**) La matrice de transfert globale, $\boldsymbol{G}(z) = \boldsymbol{H}(z)\boldsymbol{C}(z)$, vérifie $\boldsymbol{G}(z)\boldsymbol{G}^{\mathrm{H}}(1/z^*) = \boldsymbol{I}$, où $\boldsymbol{I}$ est la matrice identité $N \times N$; càd, la matrice $\boldsymbol{G}(z)$ est *para-unitaire*.

Compte-tenu de ces hypothèses, les signaux seront dits séparés si le système global $\boldsymbol{G}(z)$ s'écrit $\boldsymbol{G}(z) = \boldsymbol{\Lambda}(z)\boldsymbol{P}$, où $\boldsymbol{\Lambda}(z) = \mathbf{Diag}(e^{\jmath\psi_1}\, z^{-\alpha_1}, \ldots, e^{\jmath\psi_1}\, z^{-\alpha_N})$, où $\alpha_k \in \mathbb{Z}^N$, $1 \leq k \leq N$, et où $\boldsymbol{P}$ est une permutation.

Remarques. Si les sources ne sont plus i.i.d. mais des processus linéaires plus généraux, notre approche reste valable. Il suffit en effet de les supposer i.i.d. (**H1**) dans un premier temps pour conduire l'égalisation, puis de les estimer dans un second temps par régression linéaire entre les sorties de l'égaliseur (*i.e.* les processus générateurs) et les observations. L'hypothèse **H4** n'est pas restrictive dans le sens où on peut toujours s'y ramener en blanchissant les observations à l'ordre 2, à l'aide de n'importe quel filtre (par exemple à phase minimale) factorisant spectralement les observations [8].

3 Contrastes

Dans cette section, nous montrons comment la séparation de mélanges convolutifs peut se ramener à la diagonalisation conjointe approchée et partielle de plusieurs matrices particulières. Ceci est rendu possible par l'introduction d'un nouveau critère de contraste adapté au problème. Pour commencer, on adopte la notation suivante:

$$\begin{aligned}&\mathsf{C}^{\boldsymbol{y}}_{p,r}[i,\boldsymbol{j},\boldsymbol{\ell}] = \\ &\mathsf{Cum}[\underbrace{y_i(n),\ldots,y_i(n)}_{p\ \text{termes}},\underbrace{y_{j_1}(n-\ell_1),\ldots,y_{j_q}(n-\ell_q)}_{q=r-p\ \text{termes}}]\end{aligned} \quad (2)$$

où $r = p+q$ désigne l'ordre du cumulant, $\boldsymbol{j} = (j_1,\ldots,j_q)$, et $\boldsymbol{\ell} = (\ell_1,\ldots,\ell_q)$. Nous avons alors les deux résultats suivants:

Proposition 1 *Soient r, p et q trois entiers tels que $r \geq 3$, $2 \leq p \leq r$ et $q = r-p$, alors la fonction*

$$\mathcal{J}_{p,r}(\boldsymbol{y}) = \sum_{i=1}^{N}\sum_{\boldsymbol{j}\in\mathsf{J}}\sum_{\boldsymbol{\ell}\in\mathsf{L}}\left|\mathsf{C}^{\boldsymbol{y}}_{p,r}[i,\boldsymbol{j},\boldsymbol{\ell}]\right|^2 \quad (3)$$

est un contraste lorsque les observations $\boldsymbol{x}(n)$, et par conséquent les sorties $\boldsymbol{y}(n)$ du filtre para-unitaire, sont standardisées.

Définissons le tenseur cumulant des observations:

$$\begin{aligned}\boldsymbol{T}_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathsf{Cum}[&x_{a_1}(n-\alpha_1),\\ &x_{a_2}(n-\alpha_2),\, x_{b_1}(n-\beta_1),\ldots,x_{b_q}(n-\beta_q)]\end{aligned} \quad (4)$$

où $\boldsymbol{a}$ et $\boldsymbol{\alpha}$ sont des vecteurs de taille 2, et $\boldsymbol{b}$ et $\boldsymbol{\beta}$ sont de taille $q = r-2$, r étant l'ordre des cumulants considérés, les composantes de $\boldsymbol{a}$ et $\boldsymbol{b}$ variant dans $\{1,...,N\}$. Si L retards sont considérés, les composantes de $\boldsymbol{\alpha}$ et $\boldsymbol{\beta}$ varient dans $\{0,...,L-1\}$.

Ce tenseur peut être rangé dans une famille de matrices de taille $NL\times NL$, notées $\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})$. Pour $(\boldsymbol{b},\boldsymbol{\beta})$ fixé, les composantes de ces matrices sont données par

$$\begin{aligned}&\boldsymbol{M}_{\eta\mu}(\boldsymbol{b},\boldsymbol{\beta}) = \boldsymbol{T}_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha},\boldsymbol{\beta}),\ \text{avec}\\ &\eta = \alpha_1 N + a_1,\ \mu = \alpha_2 N + a_2\end{aligned} \quad (5)$$

Proposition 2 *Le contraste $\mathcal{J}_{2,r}(\boldsymbol{y})$ peut être écrit comme un critère de diagonalisation conjointe d'un ensemble de N^qL^q matrices:*

$$\mathcal{J}_{2,r}(\boldsymbol{y}) = \sum_{\boldsymbol{b}}\sum_{\boldsymbol{\beta}}||\mathbf{Diag}\{\mathcal{H}^{\mathrm{H}}\,\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})\,\mathcal{H}\}||^2 \quad (6)$$

où $\mathcal{H}$ est semi-unitaire.

Nous renvoyons à [8] pour de plus amples détails, pour nous consacrer plus spécialement à l'algorithmique, qui n'y avait pas du tout été abordée.

Insistons sur le fait que notre critère diffère notablement de [3] dans la mesure où non seulement les matrices $\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})$ sont très différentes, mais aussi où la matrice cherchée, $\mathcal{H}$, n'est pas carrée unitaire, mais seulement rectangulaire semi-unitaire de taille $NL\times N$, càd vérifiant $\mathcal{H}^{\mathrm{H}}\mathcal{H} = \boldsymbol{I}$. L'algorithme est donc très différent.

4 Algorithme numérique

4.1 Présentation du problème

Particularisons le résultat aux cumulants d'ordre $r = 4$ en choisissant $p = q = 2$. Les vecteurs $\boldsymbol{a}$, $\boldsymbol{\alpha}$, $\boldsymbol{b}$, et $\boldsymbol{\beta}$ sont donc de taille 2. Les propositions de la section précédente nous enseignent alors qu'il faut trouver une matrice semi-unitaire $\mathcal{H}$ de taille $NL\times N$, qui diagonalise approximativement et conjointement des N^2L^2 matrices $\boldsymbol{M}(b_1,b_2,\beta_1,\beta_2)$, chacune de taille $NL\times NL$. L'objectif est de maximiser la somme des modules carrés des N premiers éléments diagonaux des N^2L^2 matrices.

4.2 Algorithme PAJOD

Pour ce faire, on cherche une matrice unitaire $\boldsymbol{U}$ de taille $NL\times NL$, dont $\mathcal{H}$ sera la sous-matrice extraite (les N premières colonnes). Cette matrice $\boldsymbol{U}$ est construite (de manière non unique et redondante) par l'accumulation de produits de rotations planes (matrices de Givens), comme dans l'algorithme classique de Jacobi en algèbre linéaire [10] [5]:

$$\boldsymbol{U} = \prod_{1\leq i<j\leq NL}\Theta[i,j],$$

où $\Theta[i,j]$ coïncide avec la matrice identité partout sauf en quatre positions, à savoir $\Theta[i,j]_{ii} = \Theta[i,j]_{jj} = \cos(\theta[i,j])$, et $\Theta[i,j]_{ji} = -\Theta[i,j]^*_{ij} = \sin(\theta[i,j])\,e^{\jmath\psi[i,j]}$.

Les cosinus et sinus de cette rotation, notés c et s pour alléger, doivent être choisis de façon à maximiser, successivement pour chaque couple $[i,j]$:

$$\mathcal{J}_{2,4} = \sum_{\boldsymbol{b},\boldsymbol{\beta}}\sum_{k=1}^{N}\left|\sum_{\eta,\mu=1}^{NL}\Theta^*_{\eta k}\,\Theta_{\mu k}\,M_{\eta\mu}(\boldsymbol{b},\boldsymbol{\beta})\right|^2 \quad (7)$$

Cependant, les indices $[i,j]$ ne décrivent pas tous les couples possibles de $\{1,...,NL\}^2$. En effet, comme $k\leq N$, il suffit que $i\leq N$; par ailleurs $i<j$. En conséquence, on doit distinguer deux cas, suivant que $j\leq N$ ou non.

- **Cas** $j\leq N$: c'est le cas classique [4]. On maximise la somme des deux termes diagonaux sur lesquels on a une action:

  $$\begin{aligned}\mathcal{J}_{2,4} = \textstyle\sum_{\boldsymbol{b},\boldsymbol{\beta}}&|c^2M_{ii} + cs^*M_{ji} + csM_{ij} + ss^*M_{jj}|^2\\ +&|ss^*M_{ii} - cs^*M_{ji} - csM_{ij} + c^2M_{jj}|^2\end{aligned}$$

  Dans le cas réel, ceci se ramène à la résolution d'un trinôme du second degré en $X = t-1/t = -2/\tan 2\theta$:

$$\sum_{\boldsymbol{b}\boldsymbol{\beta}}(M_{ii}-M_{jj})M_{ij}(X^2-4)+(\sum_{\boldsymbol{b}\boldsymbol{\beta}}(M_{ii}-M_{jj})^2-4M_{ij}^2)X = 0$$

- **Cas** $j > N$: dans ce cas, seul le premier terme compte, et on maximise:

$$\mathcal{J}_{2,4} = \sum_{\boldsymbol{b},\boldsymbol{\beta}} |c^2 M_{ii} + cs^* M_{ji} + csM_{ij} + ss^* M_{jj}|^2$$

Dans le cas réel, on montre que ceci se ramène à la résolution d'un polynôme de degré 4 en la variable $t = \tan\theta$.

Cet algorithme a été baptisé PAJOD pour "Partial Approximate Joint Diagonalization of matrices".

5 Performances

5.1 Exemple traité

On se limite pour l'exemple aux mélanges réels FIR de longueur $L = 3$ de $N = 2$ processus réels blancs, et on choisit d'utiliser les cumulants d'ordre $r = p + q = 4$, avec $p = 2$. Nous avons donc $N^q L^q = 36$ matrices de taille $NL = 6$, dont il faut conjointement et approximativement diagonaliser le carré principal de taille 2×2, par transformation congruente à l'aide d'une matrice orthogonale réelle 6×6, $\boldsymbol{U}$. La matrice $\mathcal{H}$ correspond aux deux premières colonnes de $\boldsymbol{U}$.

Le canal a été choisi para-unitaire, de manière à ne pas détruire inutilement la blancheur des sources, conformément à ce qui avait été annoncé dans la section 2. Il a été généré comme suit:

$$\boldsymbol{C}(z) = \boldsymbol{R}(\phi_1) \cdot \boldsymbol{Z} \cdot \boldsymbol{R}(\phi_2) \cdot \boldsymbol{Z} \cdot \boldsymbol{R}(\phi_3), \text{ où}$$
$$\boldsymbol{Z} = \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix}, \text{ et } \boldsymbol{R}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

On notera l'absence volontaire de termes de phase dans les matrices unitaires $\boldsymbol{R}(\phi)$, de façon à obtenir un canal réel. En choisissant convenablement les 3 paramètres cidessus, on peut placer les 2 zéros des quatre canaux réels scalaires de longueur $L = 3$. Ici, $\phi_1 = 52^o$, $\phi_2 = 53^o$, et $\phi_3 = 70^o$, de sorte que $C_{12}(z)$ a ses zéros dans le disque unité, et $C_{21}(z)$ en dehors; $C_{11}(z)$ et $C_{22}(z)$ ont chacun un zéro dans le disque et un en dehors. Dans notre exemple, le canal $\boldsymbol{C}(z)$ admet un inverse stable, bien que ses composantes $C_{12}(z)$ et $C_{21}(z)$ n'en admettent pas.

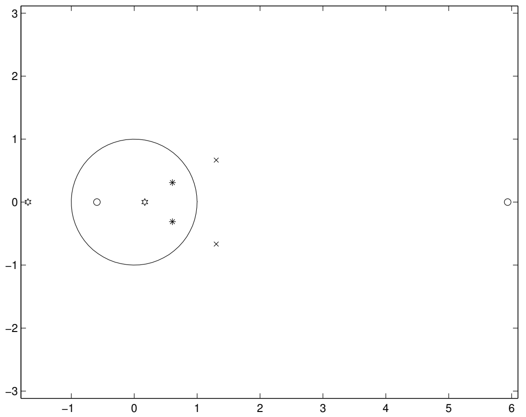

Figure 1: Zéros des quatre canaux de $\boldsymbol{C}(z)$.

5.2 Critères de performances

Une des difficultés à évaluer les performances d'un égaliseur MIMO réside dans l'indétermination inhérente au problème. En effet, l'égaliseur $\boldsymbol{H}(z)$, et par conséquent le filtre global $\boldsymbol{G}(z)$, ne peuvent être estimés qu'à une matrice $\boldsymbol{D}(z) = \boldsymbol{\Lambda}(z)\boldsymbol{P}$ près, définie dans la section 2.

5.2.1 Critère de distance

Posons $\boldsymbol{G}(z) = \sum_{n=0}^{2L-1} \boldsymbol{G}(n) z^{-n}$. On peut convenir de ranger les matrices $\boldsymbol{G}(n)$ dans un tableau $\mathcal{G}$ de taille $N \times N(2L-1)$, en les empilant les unes derrière les autres. Trouver la meilleure matrice $\boldsymbol{D}(z)$ revient à chercher dans chaque ligne de $\mathcal{G}$ l'élément de module maximum, sous la contrainte que leur numéro de colonne soit différent modulo N. Expliquons comment ceci est fait pour notre exemple où $N = 2$ et $L = 3$:

- Cas 1: chercher la colonne j_1 (resp. j_2) de l'élément de plus grand module, $\mathcal{G}_{1,j1}$ (resp. $\mathcal{G}_{2,j2}$), parmi les éléments de la ligne 1 (resp. 2) de colonnes impaires (resp. paires). Normaliser la ligne 1 (resp. 2) de $\mathcal{G}$ par $\mathcal{G}_{1,j1}$ (resp. $\mathcal{G}_{2,j2}$). Calculer la norme de Froebenius entre la matrice $\bar{\mathcal{G}}$ ainsi obtenue et la matrice $\mathcal{D}$, de même taille, contenant deux 1 aux positions $(1, j_1)$ et $(2, j_2)$ et nulle ailleurs.
- Cas 2: chercher la colonne k_1 (resp. k_2) de l'élément de plus grand module, $\mathcal{G}_{2,k1}$ (resp. $\mathcal{G}_{1,k2}$), parmi les éléments de la ligne 1 (resp. 2) de colonnes paires (resp. impaires). Répéter les mêmes opérations de normalisation et de calcul de distance que dans le cas 1.
- Choisir celui des deux cas conduisant à la distance minimale, $\epsilon(\mathcal{G}) = ||\bar{\mathcal{G}} - \mathcal{D}||$, qui correspond en réalité à:

$$\epsilon(\mathcal{G}) \stackrel{\text{def}}{=} \underset{\boldsymbol{P},\mu_1,\alpha_1,\mu_2,\alpha_2}{\text{Min}} ||\mathcal{G}\, \boldsymbol{P}\, \textbf{Diag}(\mu_1 z^{\alpha_1},\ \mu_2 z^{\alpha_2})|| \quad (8)$$

Cette procédure s'étend aisément à $N \geq 2$; on doit alors tester $N!$ cas au lieu de 2.

5.2.2 Critère d'erreur symbole

Un autre critère est celui de l'erreur sur les sources estimées. Comme précédemment, il faut s'affranchir du retard α, du facteur $\mu = e^{\jmath\psi}$ et de la permutation σ, pouvant affecter l'estimation, et minimiser $\mu\, a_{\sigma(i)}(n-\alpha) - \hat{a}_i(n)$. Ceci est fait d'une façon similaire à la section précédente, quoique plus coûteuse. En effet, sous chacune des 2 hypothèses, il faut explicitement explorer les 10 cas possibles (5 retards possibles sur chacune des lignes).

Dans le cas général, il faut calculer les taux d'erreur de $N!\, N(2L-1)$ estimateurs candidats à classer au lieu de 20, ce qui peut devenir coûteux. Dans de tels cas, on se contentera d'adopter la matrice $\mathcal{D}$ obtenue à l'aide du critère de distance.

5.3 Résultats de simulation

Les deux sources générées sont binaires, et prennent leurs valeurs dans $\{-1, 1\}$. L'égaliseur a été estimé sur des blocs de données de longueur 250 ou 400 symboles (donc très courts), bruités par du bruit blanc gaussien isotrope:

$$\boldsymbol{x}(n) = \sum_{k=0}^{L} \boldsymbol{C}(k)\boldsymbol{a}(n-k) + \rho\boldsymbol{v}(n)$$

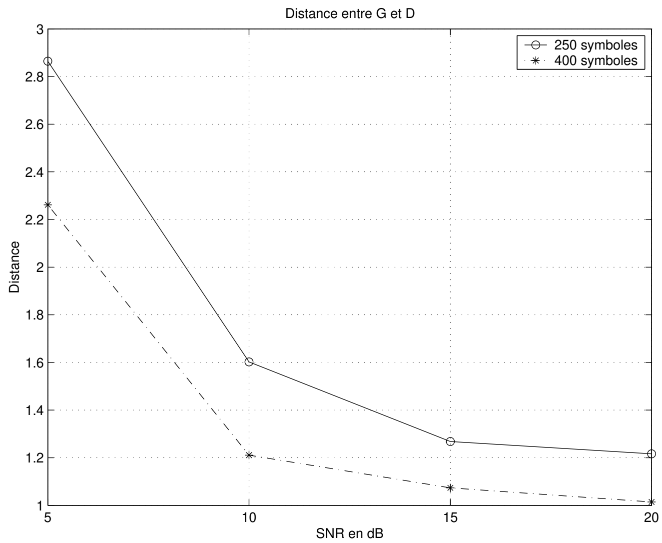

Figure 2: Erreur $\epsilon(\mathcal{G})$ sur la réponse impulsionnelle.

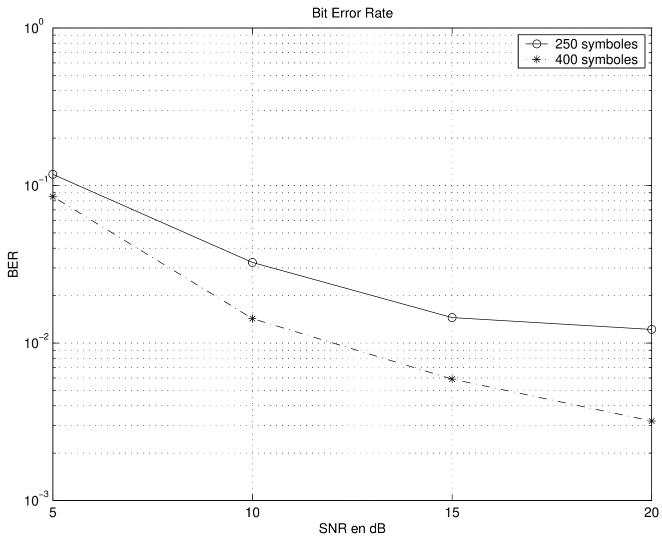

Figure 3: Taux d'Erreur Bit (BER) moyen sur les sources.

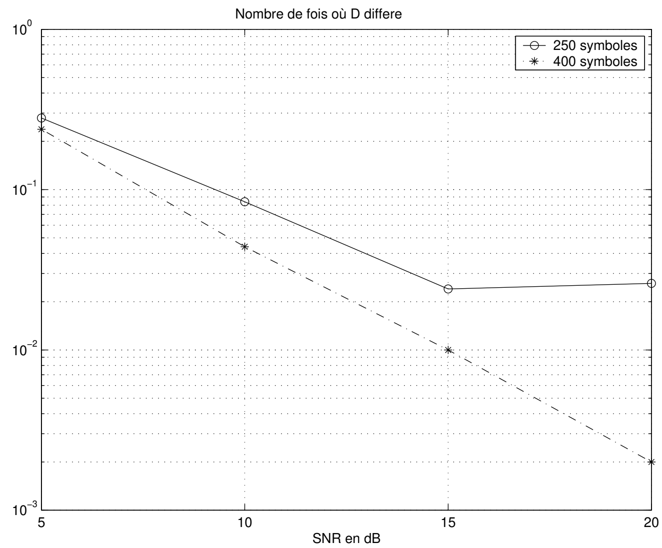

Figure 4: Taux d'erreur sur l'estimation de la matrice $\boldsymbol{D}(z) = \boldsymbol{\Lambda}(z)\boldsymbol{P}$.

Les performances ont été calculées en moyennant 500 expériences. La distance et le taux d'erreur obtenus sont représentés figures 2 et 3, en fonction du rapport signal à bruit, $-20\log_{10}\rho$.

En présence de bruit faible, les matrices $\mathcal{D}$ obtenues par les deux critères sont les mêmes. En revanche, lorsque le bruit augmente, il arrive que le critère de distance et le taux d'erreur ne conduisent pas à la même estimation de $\mathcal{D}$. C'est ce qui est montré figure 4.

Références

[1] K. ABED-MERAIM, E. MOULINES, P. LOUBATON, "Prediction error method for second-order blind identification", *IEEE Trans. Sig. Proc.*, vol. 45, no. 3, pp. 694–705, Mar. 1997.

[2] A. BENVENISTE, M. GOURSAT, "Blind equalizers", *IEEE Trans. Com.*, vol. 32, no. 8, pp. 871–883, Aug. 1984.

[3] 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.

[4] J. F. CARDOSO, A. SOULOUMIAC, "Jacobi angles for simultaneous diagonalization", *SIAM Jour. matrix Analysis*, vol. 17, no. 1, Jan. 1996.

[5] 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.

[6] P. COMON, "Contrasts for multichannel blind deconvolution", *Signal Processing Letters*, vol. 3, no. 7, pp. 209–211, July 1996.

[7] P. COMON, O. GRELLIER, "Analytical blind identification of a SISO communication channel", in *IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, PA, Usa, Aug. 14-16, 2000, pp. 206–210.

[8] P. COMON, E. MOREAU, "Blind MIMO equalization and joint-diagonalization criteria", in *ICASSP'01*, Salt Lake City, May 7-11, 2001.

[9] I. FIJALKOW, A. TOUZNI, J. R. TREICHLER, "Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity", *IEEE Trans. Sig. Proc.*, vol. 45, no. 1, pp. 56–66, Jan. 1997.

[10] G. H. GOLUB, C. F. VAN LOAN, *Matrix computations*, The John Hopkins University Press, 1989.

[11] J. G. PROAKIS, C. L. NIKIAS, "Blind equalization", in *SPIE Adaptive Signal Processing*, 1991, vol. 1565, pp. 76–88.

[12] C. SIMON, P. LOUBATON, C. VIGNAT, C. JUTTEN, G. d'URSO, "Blind source separation of convolutive mixtures by maximization of fourth-order cumulants : the non i.i.d. case", in *ICASSP*, Seattle, May 12-15 1998.

[13] A. SWAMI, G. GIANNAKIS, S. SHAMSUNDER, "Multichannel ARMA processes", *IEEE Trans. Sig. Proc.*, vol. 42, no. 4, pp. 898–913, Apr. 1994.

[14] N. THIRION, E. MOREAU, "New criteria for blind signal separation", in *IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, PA, Usa, Aug. 14-16, 2000, pp. 344–348.

[15] A. TOUZNI, I. FIJALKOW, M. LARIMORE, J. R. TREICHLER, "A globally convergent approach for blind MIMO adaptive deconvolution", in *ICASSP*, Seattle, May 12-15 1998.

[16] J. TUGNAIT, "Adaptive blind separation of convolutive mixtures of independent linear signals", *Signal Processing, Elsevier,*, vol. 73, pp. 139–152, Feb. 1999.

[17] 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.