

Estimation du nombre de sources dans les mélanges multi-entrées multi-sorties perturbés par des bruits spatialement corrélés

Pierre FABRY, Christine SERVIÈRE, Jean-Louis LACOUME

LIS, UPRESA CNRS 5083

ENSIEG, BP 46, 38402 Saint-Martin d'Hères Cedex, France

pierre.fabry@lis.inpg.fr, christine.serviere@lis.inpg.fr, jean-louis.lacoume@lis.inpg.fr

Résumé – Ce papier traite de l'estimation du nombre de sources dans des mélanges convolutifs bruités. Les traitements sont réalisés dans le domaine fréquentiel. Nous cherchons le nombre de composantes décorréliées dans chaque canal de fréquence. Les méthodes AIC et MDL habituellement utilisées supposent que les bruits sont gaussiens et spatialement blancs. Malheureusement les données réelles possèdent souvent des bruits spatialement corrélés. Lorsque les supports de corrélation des bruits sont plus faibles que ceux des sources, l'estimation reste possible. À l'aide d'observations temporellement décalées nous construisons une matrice interspectrale "débruitée". Cette matrice n'est ni hermitienne ni définie positive. Sa décomposition en valeurs singulières fournit deux matrices distinctes de vecteurs propres contenant chacune une base du sous-espace signal. Nous proposons deux critères pour estimer la dimension de ce sous-espace à partir des vecteurs propres. Des simulations confirment la robustesse de la méthode pour de forts niveaux de bruit.

Abstract – We are interested in estimating the source number in noisy convolutive mixtures. The observed signals are processed in the frequency domain. The problem is to estimate the number of uncorrelated source components at each frequency bin. The usual AIC and MDL criteria perform an hypothesis testing on the eigen values of the covariance matrix. Both methods assume that the noises are gaussian and spatially white. Unfortunately experimental data have very often correlated noises. When the noise correlation lengths are lower than the source ones, it is however possible to determine the source number. The proposed method is based on interspectral matrices computed from delayed observation vectors to cancel the noise influence. The SVD provides two matrices of eigenvectors getting each one a basis of the signal subspace. We derive two criteria to estimate the rank of the signal subspace using the eigenvectors. Simulation results are good for low SNR.

1 Introduction

Considérons un système linéaire et stationnaire à p entrées et n ($n > p$) sorties. Les entrées inconnues sont dénommées "sources". Elles sont centrées et décorréliées entre elles. Les sorties observables sont des mélanges convolutifs des sources, perturbés par des bruits additifs. Les bruits sont centrés et décorréliés des sources. Ils peuvent être corrélés entre eux. Dans le domaine fréquentiel, le nombre de composantes source mélangées n'est pas nécessairement le même à toutes les fréquences. En vue de traitements fréquentiels ultérieurs (séparation de signaux par exemple) nous souhaitons déterminer le nombre de sources présentes dans chaque canal de fréquence. Les traitements sont réalisés à partir des Transformées de Fourier Discrètes (TFD) de N échantillons. On dispose alors d'un mélange fréquentiel instantané par canal de fréquence. Notons c le nombre de composantes source présentes au canal discret f . On note $\mathbf{r}^t(f)$ le vecteur des TFD des observations temporelles sur la fenêtre $[t, \dots, t + N - 1]$. Il est modélisé par :

$$\mathbf{r}^t(f) = \underbrace{\mathbf{H}(f) \mathbf{s}^t(f)}_{\mathbf{y}^t(f)} + \mathbf{b}^t(f)$$

- $\mathbf{H}(f)$ est la matrice ($n \times c$) de mélange affectée des racines carrées des densités spectrales des sources,
- $\mathbf{s}^t(f)$ est le vecteur ($c \times 1$) des sources normalisées,
- $\mathbf{b}^t(f)$ est le vecteur ($n \times 1$) des bruits.

Dans la suite de l'exposé nous omettons la notation fréquentielle (f). La matrice spectrale des observations est définie par $\mathbf{R}_{\mathbf{r}\mathbf{r}} = \text{E} \left\{ \mathbf{r}^t \cdot (\mathbf{r}^t)^+ \right\}$ où $()^+$ désigne le transposé-conjugué. Les bruits étant centrés et décorréliés des sources, $\mathbf{R}_{\mathbf{r}\mathbf{r}} = \mathbf{R}_{\mathbf{y}\mathbf{y}} + \mathbf{R}_{\mathbf{b}\mathbf{b}}$. La matrice spectrale des mélanges $\mathbf{R}_{\mathbf{y}\mathbf{y}}$ est de rang c et s'écrit $\mathbf{R}_{\mathbf{y}\mathbf{y}} = \mathbf{H}\mathbf{H}^+$ en raison de la normalisation des sources. En introduisant la décomposition en valeurs singulières (DVS) de \mathbf{H} :

$$\mathbf{H} = \mathbf{V}\Delta^{1/2}\Pi^+, \quad \begin{cases} \mathbf{V}, \Pi \text{ unitaires } n \times c \text{ et } c \times c \\ \Delta^{1/2} = \begin{bmatrix} \sqrt{d_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{d_c} \end{bmatrix} \end{cases} \quad (1)$$

on obtient la décomposition en valeurs propres (DVP) : $\mathbf{R}_{\mathbf{y}\mathbf{y}} = \mathbf{V}\Delta\mathbf{V}^+$ où \mathbf{V} représente une base du sous-espace signal \mathcal{S} . Notre problème se ramène donc à estimer le rang de $\mathbf{R}_{\mathbf{y}\mathbf{y}}$ (où le nombre de vecteur de \mathbf{V}) à partir des observations \mathbf{r}^t .

Les méthodes AIC et MDL ([3]) couramment utilisées supposent que les bruits sont gaussiens et spatialement

blancs ($\mathbf{R}_{\mathbf{bb}} = \sigma_b \mathbf{I}_n$ où \mathbf{I}_n est l'identité de rang n). Sous cette hypothèse la DVP de $\mathbf{R}_{\mathbf{rr}}$ s'écrit :

$$\mathbf{R}_{\mathbf{rr}} = \begin{bmatrix} \mathbf{V}, \bar{\mathbf{V}} \end{bmatrix} \begin{bmatrix} \Delta + \sigma_b \mathbf{I}_c & \mathbf{0} \\ \mathbf{0} & \sigma_b \mathbf{I}_{n-c} \end{bmatrix} \begin{bmatrix} \mathbf{V}^+ \\ \bar{\mathbf{V}}^+ \end{bmatrix}$$

où $\bar{\mathbf{V}}$ est une base du sous-espace complémentaire orthogonal au sous-espace engendré par \mathbf{V} , dans l'espace des matrices ($n \times n$). Un test du Maximum de Vraisemblance est effectué sur les valeurs propres de $\mathbf{R}_{\mathbf{rr}}$. Malheureusement, les valeurs propres sont très sensibles aux perturbations [1]. Pour peu que les bruits soient spatialement corrélés, de puissances différentes, ou que la matrice spectrale ne soit pas assez bien estimée, le test d'hypothèses est mis en défaut.

Nous présentons une méthode qui accepte la présence de bruits spatialement corrélés, de lois et de fonctions de corrélation inconnues. Cette méthode suppose simplement que les supports de corrélation des bruits sont plus faibles que ceux des sources.

Nous montrons en section 2 qu'en utilisant des observations temporellement décalées il est possible d'obtenir une matrice interspectrale "débruitée" que nous dénommons "matrice spectrale dissymétrique". Sa DVS fournit une matrice diagonale de valeurs singulières et deux matrices distinctes de vecteurs propres. En théorie, seules les c premières valeurs singulières ne sont pas nulles. Les vecteurs propres qui leur sont associés à gauche et à droite constituent deux bases du sous-espace signal \mathcal{S} . En présence de perturbations (matrice spectrale mal estimée, mauvais choix du retard, bruit très fort) les écarts relatifs entre les c premières valeurs singulières changent et les dernières valeurs singulières ne sont plus nulles. En revanche les c premiers vecteurs de gauche et de droite restent fortement représentatifs de \mathcal{S} . Il est donc préférable d'exploiter les vecteurs propres plutôt que les valeurs singulières pour déterminer la dimension de \mathcal{S} .

Aux sections 3 et 4 nous proposons deux critères pour estimer la dimension de \mathcal{S} à partir des vecteurs propres de la matrice spectrale dissymétrique.

Enfin nous illustrons les performances de ces méthodes par des résultats de simulations en fonction du Rapport Signal à Bruit.

2 Matrice spectrale dissymétrique

Définissons la matrice spectrale dissymétrique des observations par :

$$\mathbf{R}_{\mathbf{rr}}^\tau = E \{ \mathbf{r}^t \cdot (\mathbf{r}^{t+\tau})^+ \} = \mathbf{R}_{\mathbf{yy}}^\tau + \mathbf{R}_{\mathbf{bb}}^\tau$$

où \mathbf{r}^t et $\mathbf{r}^{t+\tau}$ sont les TFD des observations temporelles sur les fenêtres débutant respectivement aux indices t et $t + \tau$. Notons τ_s le plus petit support de corrélation des sources et τ_B le plus grand support de corrélation des bruits. Si l'on peut choisir un retard τ tel que $N + \tau_B < \tau \leq \tau_s$ alors \mathbf{r}^t et $\mathbf{r}^{t+\tau}$ sont issus d'échantillons de bruit décorrelés et d'échantillons de signal corrélés. Par conséquent $\mathbf{R}_{\mathbf{bb}}^\tau$ est nulle et :

$$\mathbf{R}_{\mathbf{rr}}^\tau = \mathbf{R}_{\mathbf{yy}}^\tau = \mathbf{H} \mathbf{R}_{\mathbf{ss}}^\tau \mathbf{H}^+ \quad (2)$$

En raison de la normalisation et de l'indépendance des sources, $\mathbf{R}_{\mathbf{ss}}^\tau$ est une matrice diagonale d'exponentielles complexes, encore notée θ :

$$\mathbf{R}_{\mathbf{ss}}^\tau = E \{ \mathbf{s}^t \cdot (\mathbf{s}^{t+\tau})^+ \} = \begin{bmatrix} e^{j\theta_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & e^{j\theta_c} \end{bmatrix} = \theta \quad (3)$$

En remplaçant (1) et (3) dans (2) on obtient :

$$\mathbf{R}_{\mathbf{rr}}^\tau = \underbrace{\mathbf{V}}_{n \times c} \underbrace{\Delta^{1/2} \Pi^+ \theta \Pi \Delta^{1/2}}_{\mathbf{A}} \underbrace{\mathbf{V}^+}_{c \times n} \quad (4)$$

La matrice \mathbf{A} , de rang c , admet aussi une DVS de la forme

$$\mathbf{A} = \mathbf{Y} \Lambda^{1/2} \mathbf{Z}^+, \quad \begin{cases} \mathbf{Y}, \mathbf{Z} \text{ unitaires } c \times c \\ \Lambda^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_c} \end{bmatrix} \end{cases} \quad (5)$$

où les valeurs singulières vérifient $\lambda_1 \geq \dots \geq \lambda_c > 0$. En remplaçant (5) dans (4) on obtient une écriture réduite de la DVS de $\mathbf{R}_{\mathbf{rr}}^\tau$:

$$\mathbf{R}_{\mathbf{rr}}^\tau = \mathbf{E}_c \Lambda^{1/2} \mathbf{F}_c^+ \quad \text{avec} \quad \begin{cases} \mathbf{E}_c = \mathbf{V} \mathbf{Y} \\ \mathbf{F}_c = \mathbf{V} \mathbf{Z} \end{cases} \quad (6)$$

On constate que \mathbf{E}_c et \mathbf{F}_c sont liés à \mathbf{V} par une transformation orthogonale. Elles constituent deux bases distinctes de \mathcal{S} . Les transformations orthogonales sont induites par la matrice de phases θ qui, dans le cas général, n'est pas proportionnelle à l'identité. L'écriture complète de la DVS de $\mathbf{R}_{\mathbf{rr}}^\tau$ s'écrit

$$\mathbf{R}_{\mathbf{rr}}^\tau = \underbrace{[\mathbf{E}_c, \bar{\mathbf{E}}_c]}_{\mathbf{E}} \begin{bmatrix} \Lambda^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \underbrace{[\mathbf{F}_c, \bar{\mathbf{F}}_c]^+}_{\mathbf{F}} \quad (7)$$

où \mathbf{E} et \mathbf{F} de dimension $n \times n$ sont les matrices unitaires de vecteurs propres contenant chacune :

- une base orthonormée de \mathcal{S} (respectivement \mathbf{E}_c et \mathbf{F}_c de dimension $n \times c$),
- une base orthonormée du complémentaire \mathcal{S}^\perp (respectivement $\bar{\mathbf{E}}_c$ et $\bar{\mathbf{F}}_c$ de dimension $n \times (n - c)$).

Pour estimer c nous proposons tout d'abord un critère heuristique qui analyse l'ensemble des produits scalaires entre les vecteurs de \mathbf{E} et \mathbf{F} . Ensuite nous envisageons un critère algébrique qui analyse l'ensemble des distances entre les sous-espaces.

3 Critère heuristique

Si nous réalisons les produits scalaires entre les vecteurs propres à gauche et à droite de $\mathbf{R}_{\mathbf{rr}}^\tau$, nous obtenons la matrice unitaire :

$$\mathbf{A} = \mathbf{E}^+ \mathbf{F} = \begin{bmatrix} \mathbf{E}_c^+ \mathbf{F}_c & \mathbf{E}_c^+ \bar{\mathbf{F}}_c \\ \bar{\mathbf{E}}_c^+ \mathbf{F}_c & \bar{\mathbf{E}}_c^+ \bar{\mathbf{F}}_c \end{bmatrix}$$

Les sous-matrices $\mathbf{E}_c^+ \bar{\mathbf{F}}_c$ et $\bar{\mathbf{E}}_c^+ \mathbf{F}_c$ sont nulles car le sous-espace engendré par \mathbf{E}_c et \mathbf{F}_c est orthogonal à celui engendré par $\bar{\mathbf{E}}_c$ et $\bar{\mathbf{F}}_c$. Les sous-matrices $\mathbf{E}_c^+ \mathbf{F}_c = \mathbf{Y}^+ \mathbf{Z}$ et

$\overline{\mathbf{E}}_c^+ \overline{\mathbf{F}}_c$ sont unitaires. Nous les dénommerons respectivement \mathbf{A}_1 et \mathbf{A}_2 . La matrice \mathbf{A} possède donc une structure particulière :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix} \begin{matrix} c \\ n-c \\ c & n-c \end{matrix}$$

La relation (4) nous indique que lorsque θ est proportionnelle à l'identité, les c premiers vecteurs propres à gauche et à droite de $\mathbf{R}_{\mathbf{r}\mathbf{r}}$ sont identiques. Par conséquent \mathbf{A}_1 est l'identité de rang c . Cette situation (où une situation proche) peut se produire pour certaines valeurs du retard τ . La matrice \mathbf{A}_2 quant à elle est généralement pleine. En effet, les vecteurs colonne de $\overline{\mathbf{E}}_c$ n'ont aucune raison d'être identiques à ceux de $\overline{\mathbf{F}}_c$ puisqu'ils sont associés à des valeurs singulières nulles et qu'ils sont, de ce fait, très sensibles aux perturbations [1]. La sous-matrice \mathbf{A}_2 possède une structure qui la rend plus facilement détectable que \mathbf{A}_1 et qui permet de déterminer les dimensions de \mathcal{S} et \mathcal{S}^\perp . La figure 1 représente les modules des produits scalaires dans le cas de 2 sources et de 6 capteurs sous un Rapport Signal à Bruit de -5 dB.

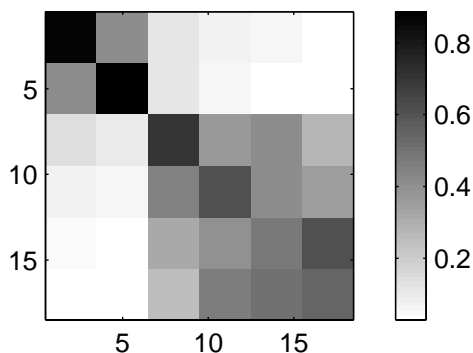


FIG. 1: $|\mathbf{A}|$, 2 sources, 6 capteurs, RSB=-5dB

Ce critère visuel nous autorise à traiter $|\mathbf{A}| = (|a_{ij}|)$ comme une image pour détecter les contours gauche et supérieur de $|\mathbf{A}_2|$. Nous commençons par dilater $|\mathbf{A}|$ d'un facteur 3 en lignes et en colonnes de façon à pouvoir appliquer des filtres RIF bidimensionnels. Un filtre Laplacien 3×3 permet de rehausser les contrastes (figure 2).

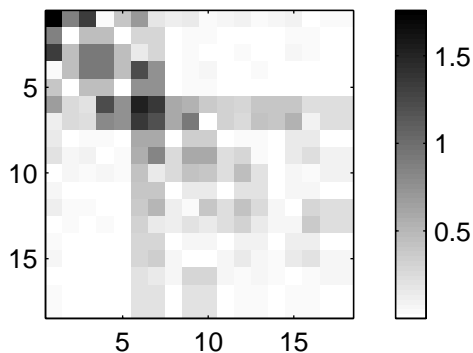


FIG. 2: résultat du filtrage Laplacien

Il est suivi d'un seuillage où les pixels positifs sont remplacés par la valeur 1 et les autres par la valeur 0. On

constate alors (figure 3) que les contours gauche et supérieur de $|\mathbf{A}_2|$ (indice $3 \cdot c = 6$) ont été remplacés par une suite de 1 et définissent la séparation entre \mathcal{S} de dimension c et \mathcal{S}^\perp de dimension $n - c$.

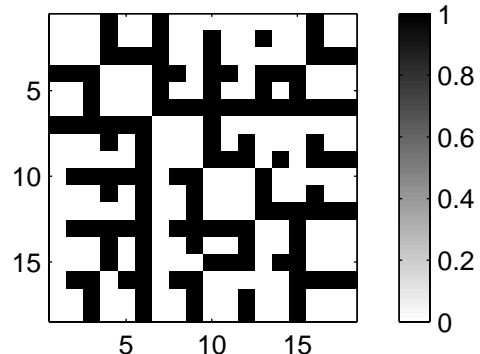


FIG. 3: résultat du seuillage

Le cas où $|\mathbf{A}|$ est une matrice pleine correspond soit à l'absence de sources soit à la présence de n sources. Pour lever l'ambiguïté nous supposons que le nombre de capteurs est toujours strictement supérieur au nombre de sources. Le critère heuristique permet donc de déterminer la présence de 0 à $n - 1$ sources.

Pour améliorer la stabilité de la méthode nous utilisons deux matrices spectrales dissymétriques (retards τ_1 et τ_2) de façon à disposer de 4 matrices de vecteurs propres et à obtenir $|\mathbf{A}|$ par la moyenne de 12 produits scalaires. Ceci permet de s'assurer que les contours de $|\mathbf{A}_2|$ ne possèdent pas de "trous". La détection est améliorée.

4 Distance entre sous-espaces

Dans le but de définir un critère algébrique nous considérons toutes les partitions d'indice k de \mathbf{E} et \mathbf{F} , pour k variant de 1 à $n - 1$:

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_k & \overline{\mathbf{E}}_k \\ n \times k & n \times (n-k) \end{bmatrix} ; \quad \mathbf{F} = \begin{bmatrix} \mathbf{F}_k & \overline{\mathbf{F}}_k \\ n \times k & n \times (n-k) \end{bmatrix}$$

La seule partition pour laquelle \mathbf{E}_k et \mathbf{F}_k engendrent le même sous-espace est celle d'indice k égal à c . Il suffit donc de mesurer la distance entre les sous-espaces engendrés par \mathbf{E}_k et \mathbf{F}_k , notés respectivement $\mathfrak{S}(\mathbf{E}_k)$ et $\mathfrak{S}(\mathbf{F}_k)$. D'après [1], cette distance est définie par la norme-2 matricielle de l'écart entre les projecteurs orthogonaux sur $\mathfrak{S}(\mathbf{E}_k)$ et $\mathfrak{S}(\mathbf{F}_k)$ (8). En effet, le projecteur orthogonal sur un espace vectoriel est unique; il est défini par $\mathbf{X}\mathbf{X}^+$ ou \mathbf{X} est une base orthonormée de cet espace. La distance entre $\mathfrak{S}(\mathbf{E}_k)$ et $\mathfrak{S}(\mathbf{F}_k)$ est donc définie par :

$$d_k = \|\mathbf{E}_k \mathbf{E}_k^+ - \mathbf{F}_k \mathbf{F}_k^+\|_2 \quad (8)$$

Du point de vue numérique, l'estimation de c est obtenue par

$$\hat{c} = \arg \left[\min_k (d_k) \right]$$

Ce critère permet de déterminer entre 1 et $n - 1$ sources, car d'une part il existe toujours un minimum et d'autre part pour $k = n$ la distance est toujours nulle ($\mathbf{E}_n = \mathbf{E}$ et $\mathbf{F}_n = \mathbf{F}$ sont deux bases de l'espace complet).

5 Résultats de simulation

Deux fréquences pures proches filtrées par des filtres AR1 sont reçues sur six capteurs. Les observations sont perturbées par des bruits gaussiens temporellement blancs. Nous avons simulé le cas de bruits spatialement blancs pour pouvoir comparer les nouvelles méthodes aux critères AIC et MDL et le cas de bruits spatialement corrélés. Les TFD sont calculées sur des blocs de 64 échantillons fournissant une résolution de $\frac{1}{64} \approx 0.016$ en fréquence réduite. Les fréquences réduites des sources étant $f_1 = 0.10$ et $f_2 = 0.11$, une analyse spectrale à cette résolution ne permet pas de les dissocier. Les traitements sont réalisés au canal $f = 0.11$ pour différents RSB. Les matrices interspectrales sont estimées sur 600 blocs. Le retard τ est de 67 échantillons. Nous présentons des probabilités de succès estimées à l'aide de 100 réalisations pour chaque niveau de bruit.

La figure 4 représente des probabilités de succès en fonction du RSB dans le cas de bruits spatialement blancs. La méthode heuristique qui utilise des traitements d'image est notée *ti* et la méthode par distance minimale entre sous-espaces qui utilise les projecteurs orthogonaux est notée *po*. On constate la supériorité de la méthode par distance entre sous-espaces.

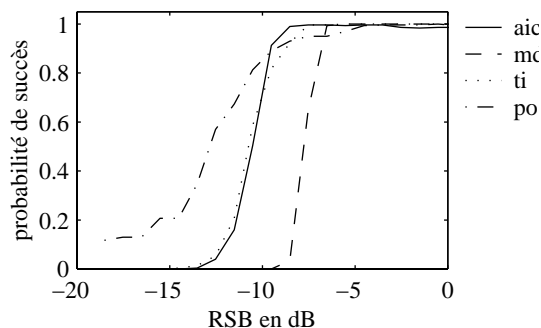


FIG. 4: probabilité de succès en présence de bruits spatialement blancs

La figure 5 représente des probabilités de succès en fonction du RSB dans le cas de bruits spatialement corrélés. On constate que les performances de nos deux critères sont peu affectées par la présence de bruits spatialement corrélés. Les méthodes AIC et MDL (non représentées) ne fonctionnent absolument pas, dans ce contexte.

6 Conclusions

Dans ce papier nous nous sommes intéressés aux mélanges MIMO (Multiple Input Multiple Output) perturbés par des bruits spatialement corrélés. Le problème était de déterminer le nombre de composantes mélangées par bande de fréquence. En supposant que les supports de corrélation des sources sont plus grands que ceux des bruits, nous développons une méthode qui utilise un retard pour supprimer l'influence du bruit. Aucune hypothèse n'est nécessaire, concernant les lois des bruits. Contrairement aux méthodes classiques qui exploitent les valeurs propres

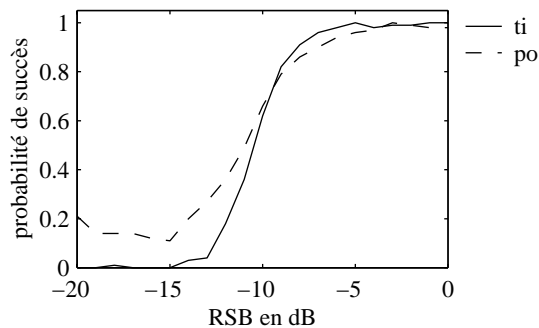


FIG. 5: probabilité de succès en présence de bruits spatialement corrélés

d'une matrice de covariance, nos critères exploitent les vecteurs propres d'une matrice de covariance dissymétrique. Nous avons proposé deux critères pour retrouver la dimension du sous-espace signal à partir des vecteurs propres. Nos critères sont robustes aux corrélations spatiales des bruits et permettent d'envisager de faibles valeurs de RSB. Dans le cas de bruits spatialement blancs, nos critères donnent de meilleurs résultats que les méthodes classiques (AIC et MDL). Ceci s'explique par une meilleure stabilité des vecteurs propres aux perturbations.

Références

- [1] G.H. Golub and F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 1983.
- [2] V.L. Koliadin. Determination of the signal-subspace dimension : A nonparametric approach based on the eigenvectors analysis. *Signal Processing VII*, pages 268–271, 1994.
- [3] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, Vol. 6(No. 2):461–464, 1978.
- [4] P. Stoica and M. Cedervall. Detection tests for array processing in unknown correlated noise fields. *IEEE transactions on signal processing*, 45(9):2351–2362, September 1997.
- [5] Q.T. Zhang and Kon Max Wong. Information theoretic criteria for the determination of the number of signals in spatially correlated noise. *IEEE Transactions on signal processing*, 41(4):1652–1663, April 1993.