

# Projet ANR-12-MONU-0003 DIONISOS.

## Rapport D3.1-2: Problèmes liés aux sources large bande.

### 1 Introduction.

Ce rapport est le premier document concernant les travaux consacrés au WP3 intitulé "Problèmes liés aux sources large bande" alors que un premier rapport était prévu à  $t_0 + 12$  dans le document de soumission. Cela est dû aux difficultés techniques rencontrées lors de l'étude de la localisation des valeurs propres de la matrice de covariance spatio-temporelle empirique d'un bruit blanc gaussien dans certains régimes asymptotiques. Nous avons en particulier cherché à obtenir un résultat que nous espérons optimal, et sa preuve s'est avérée difficile à mettre en place. En pratique, un article a été finalement soumis en Mai 2014:

P. Loubaton, "On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices", arXiv:1405.2006.

A la suite de ce résultat, nous aurions pu nous intéresser immédiatement au problème de la détection d'une source large bande, mais nous avons préféré nous consacrer à l'analyse des performances d'algorithmes de localisation de sources bande étroite basés sur la technique dite du "spatial smoothing", dont l'intérêt est de permettre l'utilisation de méthodes sous-espace dans le cas où le nombre d'observations est très nettement inférieur au nombre de capteurs. Les applications à la détection de source large bande seront développées après la finalisation de ce travail qui a donné lieu à la soumission d'un article à la conférence ICASSP 2015:

G.T. Pham, P. Loubaton, P. Vallet, "Performance analysis of spatial smoothing schemes in the context of large arrays"

qui est disponible sur le site web du projet.

Dans le présent rapport, nous allons résumer les résultats de ces 2 articles et tracer quelques perspectives à court terme pour la suite du projet.

### 2 Positionnement des travaux.

Rappelons que le WP3 est consacré à des problèmes de détection / estimation multi-capteurs mettant en jeu des signaux large bande. En présence d'une source large bande, le signal de dimension  $M$   $(\mathbf{y}_n)_{n \in \mathbb{Z}}$  recueilli sur le réseau de capteurs à  $M$  éléments peut s'écrire sous la forme

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{a}_p s_{n-p} + \mathbf{v}_n = \mathbf{x}_n + \mathbf{v}_n \quad (2.1)$$

où  $(s_n)_{n \in \mathbb{Z}}$  représente un signal scalaire déterministe non observable,  $(\mathbf{a}_p)_{p=0, \dots, P-1}$  sont des vecteurs déterministes inconnus de dimension  $M$  et où  $(\mathbf{v}_n)_{n \in \mathbb{Z}}$  est un bruit gaussien complexe, blanc temporellement et spatialement de variance  $\sigma^2$ . Le signal  $\mathbf{x}_n$  représente le signal utile sur lequel il est souhaitable d'obtenir des informations à partir des observations  $(\mathbf{y}_n)_{n=1, \dots, N}$ . On peut par exemple s'intéresser aux problèmes suivants:

- Détecter la présence ou l'absence de  $(\mathbf{x}_n)$  dans l'observation, ce qui est équivalent à un problème de détection.

- Estimer les vecteurs  $(\mathbf{a}_p)_{p=0,\dots,P-1}$ .
- Estimer la sequence  $(s_n)$ .

(voir par exemple [15], [10], [16], [1] pour plus de précisions). Un bon nombre d'algorithmes de détection / estimation mettant en jeu le signal  $\mathbf{y}$  utilisent les valeurs propres et les vecteurs propres de la matrice  $\frac{\mathbf{Y}_L \mathbf{Y}_L^*}{N}$  où  $\mathbf{Y}_L$  est la matrice Hankel par blocs  $ML \times N$  définie par

$$\mathbf{Y}_L = \begin{pmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \cdots & \mathbf{y}_{N-L+1} \\ \mathbf{y}_2 & \mathbf{y}_3 & \ddots & \ddots & \mathbf{y}_{N-L+2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{y}_L & \mathbf{y}_{L+1} & \cdots & \cdots & \mathbf{y}_N \end{pmatrix}$$

La matrice  $\mathbf{Y}_L$  est la somme de la matrice déterministe  $\mathbf{X}_L$  et de la matrice aléatoire  $\mathbf{V}_L$ , toutes deux définies comme  $\mathbf{Y}_L$ . En principe, l'entier  $L$  est choisi plus grand que  $P$  car, dans ce cas, la matrice  $ML \times ML$   $\frac{\mathbf{X}_L \mathbf{X}_L^*}{N}$  est de rang déficient quand  $ML < N$  dès que  $M \geq 2$ , ce qui s'avère utile dans le cadre des problèmes de détection / estimation listés plus haut.

Les algorithmes classiques de détection / estimation sont bien compris quand  $ML$  est très petit par rapport à  $N$ , ce que l'on modélise par le régime asymptotique  $ML$  fixe et  $N \rightarrow +\infty$ . La raison principale est que, dans ce contexte, la loi des grands nombres aboutit à

$$\left\| \frac{\mathbf{Y}_L \mathbf{Y}_L^*}{N} - \left( \frac{\mathbf{X}_L \mathbf{X}_L^*}{N} + \sigma^2 \mathbf{I}_{ML} \right) \right\| \rightarrow 0$$

Cependant, il arrive souvent que  $ML$  et  $N$  soient du même ordre de grandeur. Il est donc intéressant de s'intéresser aux régimes asymptotiques dans lesquels

$$M \rightarrow +\infty, N \rightarrow +\infty, \frac{ML}{N} \rightarrow c \quad (2.2)$$

avec  $c > 0$  et d'étudier le comportement de  $\frac{\mathbf{Y}_L \mathbf{Y}_L^*}{N}$  dans ce contexte. Il convient de noter que la condition (2.2) couvre différents types de régimes:  $L$  fixe et  $M$  et  $N$  du même ordre de grandeur, tandis que  $L$  peut tendre vers l'infini beaucoup plus vite que  $M$ . Il convient de noter que nous n'envisageons par le cas où  $M$  est fixe, ou de façon équivalente, le cas où  $L$  et  $N$  tendent vers l'infini au même rythme. Dans ce contexte, les techniques à utiliser sont complètement différentes et nettement moins matures sur le plan mathématique.

Afin d'étudier les algorithmes de détection et d'estimation relatifs aux modèles large bande dans le régime (2.2), nous étudions les éléments propres de la matrice  $\frac{\mathbf{Y}_L \mathbf{Y}_L^*}{N}$ . Pour cela, notre démarche a été de considérer dans un premier temps le cas où le signal utile est absent dans l'espoir de pouvoir utiliser une démarche proche de celle qui est utilisée dans [5] pour traiter le cas où une source est présente.

### 3 Localisation des valeurs propres de la matrice de covariance spatio-temporelle empirique en absence de signal.

Dans ce cadre, la matrice  $\mathbf{Y}_L$  est réduite à  $\mathbf{V}_L$ , de sorte qu'il convient d'étudier les valeurs propres de  $\frac{\mathbf{V}_L \mathbf{V}_L^*}{N}$ . Afin de permettre au lecteur de mieux connecter ce rapport avec l'article [8], nous utilisons les notations de [8]. Dans [8], la matrice étudiée, appelée  $\mathbf{W}_N$ , est obtenue à partir de  $\frac{1}{\sqrt{N}} \mathbf{V}_L$  en permutant ses lignes et ses colonnes. Par ailleurs, le nombre d'observations  $N$  est changé dans [8] en  $N + L - 1$  afin d'y simplifier les notations. Comme nous n'étudions que le cas où  $\frac{L}{N} \rightarrow 0$ , ceci est sans conséquence dans l'interprétation des résultats de [8].

Dans [8], la matrice  $\mathbf{W}_N$  est obtenue à partir de variables aléatoires gaussiennes complexes scalaires  $(w_{m,n})_{m=1,\dots,M,n=1,\dots,N+L-1}$  indépendantes identiquement distribuées vérifiant  $\mathbb{E}|w_{m,n}|^2 = \frac{\sigma^2}{N}$  et  $\mathbb{E}(w_{m,n}^2) = 0$ . Si on définit les matrices  $L \times N$  de Hankel  $(\mathbf{W}_N^{(m)})_{m=1,\dots,M}$  dont les éléments sont donnés par

$$\left( \mathbf{W}_N^{(m)} \right)_{i,j} = w_{m,i+j-1}, \quad 1 \leq i \leq L, 1 \leq j \leq N \quad (3.3)$$

alors la matrice  $\mathbf{W}_N$  représente la matrice  $ML \times N$  donnée par

$$\mathbf{W}_N = \begin{pmatrix} \mathbf{W}_N^{(1)} \\ \mathbf{W}_N^{(2)} \\ \vdots \\ \mathbf{W}_N^{(M)} \end{pmatrix} \quad (3.4)$$

Dans [8], nous démontrons les résultats suivants:

- La distribution empirique des valeurs propres de  $\mathbf{W}_N \mathbf{W}_N^*$  a le même comportement que si les éléments de  $\mathbf{W}_N$  étaient i.i.d., c'est-à-dire qu'elle converge presque sûrement vers la distribution de Marcenko-Pastur  $\mu_{\sigma^2, c}$  de paramètres  $(\sigma^2, c)$ .
- Par ailleurs, si  $L = \mathcal{O}(N^\alpha)$  avec  $\alpha < 2/3$ , alors, pour tout  $\epsilon > 0$  aussi petit que l'on veut, presque sûrement, pour tout  $N$  assez grand, les  $\min(ML, N)$  valeurs propres non nulles de  $\mathbf{W}_N \mathbf{W}_N^*$  sont localisées dans l'intervalle  $[\sigma^2(1 - \sqrt{c})^2 - \epsilon, \sigma^2(1 + \sqrt{c})^2 + \epsilon]$

Rappelons que la distribution de Marcenko-Pastur  $\mu_{\sigma^2, c}$  ([9]) est la mesure de probabilité donnée par

$$d\mu_{\sigma^2, c}(x) = \delta_0[1 - c^{-1}]_+ + \frac{\sqrt{(x - x^-)(x^+ - x)}}{2\sigma^2 c \pi x} \mathbb{1}_{[x^-, x^+]}(x) dx$$

où  $x^- = \sigma^2(1 - \sqrt{c})^2$  and  $x^+ = \sigma^2(1 + \sqrt{c})^2$ , et que de façon informelle, la convergence de la distribution empirique des valeurs de  $\mathbf{W}_N \mathbf{W}_N^*$  vers  $\mu_{\sigma^2, c}$  signifie que l'histogramme des valeurs propres de toute réalisation de  $\mathbf{W}_N \mathbf{W}_N^*$  tend à se rapprocher du graphe de la densité de probabilité de  $\mu_{\sigma^2, c}$ . Ce résultat est valable dès que  $M$  tend vers l'infini, même si cette convergence est très lente par rapport à celle de  $N$ . Cette convergence n'est pas très étonnante au vu de publications récentes consacrées aux valeurs propres de certaines grandes matrices aléatoires structurées, par exemple des matrices hermitiennes dont les blocs sont des matrices de Toeplitz ([3]) indépendantes entre elles et construites à partir de variables aléatoires i.i.d.. La localisation des valeurs propres non nulles de  $\mathbf{W}_N \mathbf{W}_N^*$  au voisinage du support de la densité de probabilité de  $\mu_{\sigma^2, c}$  est également un résultat attendu, mais il semble qu'il ne soit pas valable si  $L$  tend vers l'infini trop vite par rapport à  $M$ . La condition  $\alpha < 2/3$ , qui correspond à  $\frac{L}{M^2} \rightarrow 0$  apparait en effet au cours de diverses étapes indépendantes de la preuve de [8], et il semble qu'elle ne puisse pas se généraliser à des valeurs de  $\alpha$  plus grandes que  $2/3$ . En tout état de cause, les valeurs de  $M, N, L$  susceptibles d'être modélisées par la condition  $L = \mathcal{O}(N^\alpha)$  avec  $\alpha < 2/3$  nous semblent suffisamment diverses pour couvrir les applications usuelles dans le cas où le nombre de capteurs disponibles est grand. En pratique, pour de telles valeurs de  $M, N, L$ , le résultat de localisation que nous avons établi signifie que la probabilité que l'une des valeurs propres de  $\mathbf{W}_N \mathbf{W}_N^*$  s'échappe du support de la distribution de Marcenko-Pastur est faible. Nous préférons ne pas évoquer les aspects techniques de la preuve des résultats annoncés, et nous contentons de mentionner que notre approche est basée sur les techniques développées dans [6] et [12] dans le cas  $L = 1$ .

Nous donnons quelques exemples numériques permettant d'illustrer la localisation des valeurs propres. La figure 1 représente les valeurs propres d'une réalisation de  $\mathbf{W}_N \mathbf{W}_N^*$  dans le cas où  $M = 20, L = 5, N = 200, c = 1/2, \sigma^2 = 1$ . Dans ce scénario, le rapport  $L/M^2$  est égal à  $L/M^2 = 0.0125$ , de sorte que les paramètres modélisent bien une situation dans laquelle  $L/M^2 \simeq 0$ . Avec ces paramètres, il est difficile de trouver une réalisation de  $\mathbf{W}_N \mathbf{W}_N^*$  dans laquelle l'une des valeurs propres sort du support de la distribution de Marcenko-Pastur.

Dans la figure 2, les paramètres valent  $M = 20, L = 60, N = 2400, c = 1/2, \sigma^2 = 1$ , ce qui correspond à  $L/M^2 = 0.15$  qui n'est pas précisément négligeable. Dans la réalisation choisie, toutes les valeurs propres restent dans le support, mais contrairement au contexte de la figure 3, il est difficile de trouver une telle réalisation. Ainsi, la figure 3, correspondant à un autre tirage de  $\mathbf{W}_N \mathbf{W}_N^*$ , fait apparaître des valeurs propres s'échappant du support de la distribution de Marcenko-Pastur.

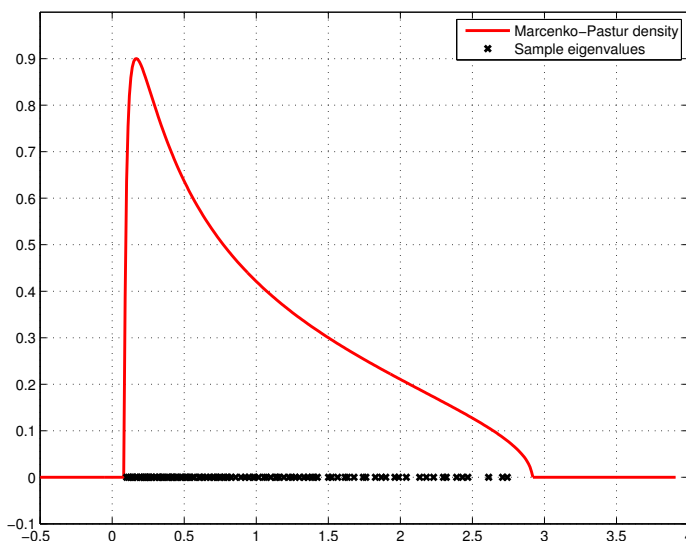


Figure 1:  $M = 20, L = 5, L/M^2 = 0.0125, N = 200, c = 1/2, \sigma^2 = 1$

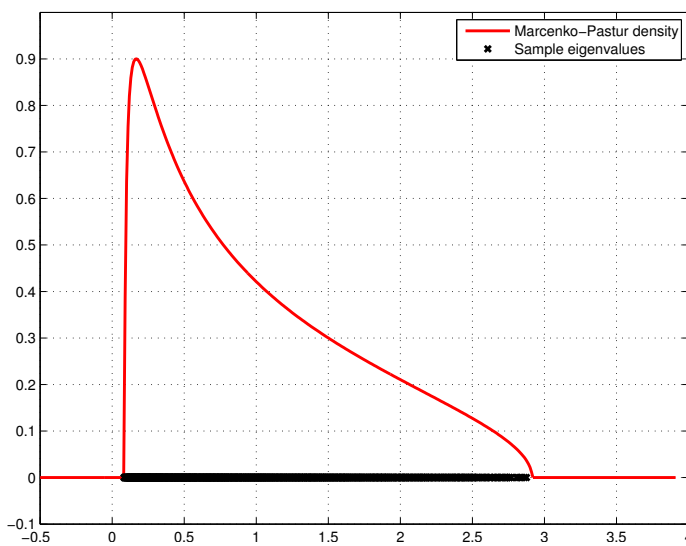


Figure 2:  $M = 20, L/M^2 = 0.15, L = 60, N = 2400, c = 1/2, \sigma^2 = 1$

Il convient de mentionner que si  $M$  ne tend pas l'infini, les résultats que nous avons établi dans [8] ne sont plus valables. Dans ce contexte, la distribution des valeurs propres de  $\mathbf{W}_N \mathbf{W}_N^*$  ne converge pas vers la distribution de Marcenko-Pastur, mais vers une distribution à support non compact plus ou moins caractérisée par ses moments, voir par exemple [2] qui a considéré le cas  $M = 1$ . Le comportement de la plus grande valeur propre de  $\mathbf{W}_N \mathbf{W}_N^*$  n'est pas étudié dans [2], mais, au vu des résultats de [13] (qui concernent un modèle différent), il semble probable qu'elle tende vers  $\infty$  à un rythme qui devrait être  $\log N$ . En pratique, la rupture entre le régime  $M \rightarrow +\infty$  et celui dans lequel  $M$  reste fixe alors que  $L$  et  $N$  tendent vers l'infini au même rythme est progressive, et se matérialise par des histogrammes de valeurs propres de  $\mathbf{W}_N \mathbf{W}_N^*$  s'ajustant de moins en moins bien à la densité de probabilité de la distribution de Marcenko-Pastur.

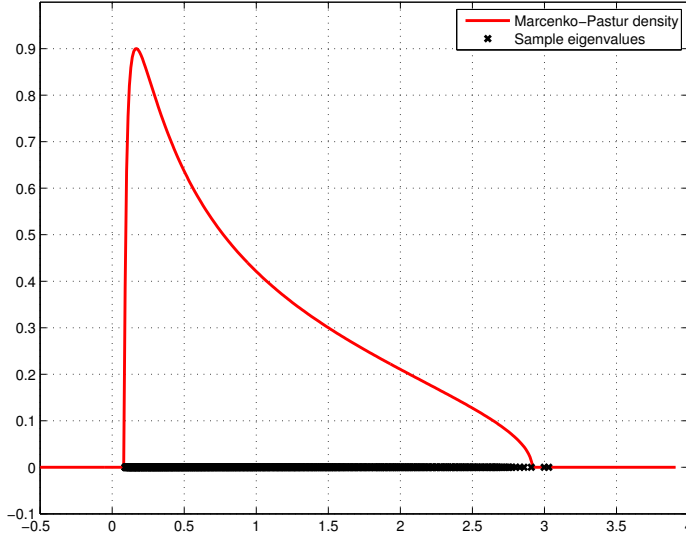


Figure 3:  $M = 20, L/M^2 = 0.15, L = 60, N = 2400, c = 1/2, \sigma^2 = 1$

## 4 Application à l'analyse des performances des techniques de localisation de sources bande étroite utilisant la technique du lissage spatial.

Dans un premier temps, nous avons choisi d'appliquer le résultat précédent dans un contexte de source bande étroite mais dans lequel les matrices de covariances empiriques ont des structures qui sont proches de celles qui viennent d'être évoquées. Nous supposons que  $K$  sources bande étroite situées à l'infini se propagent, et que les signaux qu'elles produisent sont reçus par un réseau linéaire à capteurs équidistants de  $M$  antennes, avec  $K < M$ . Dans ce contexte, le signal de dimension  $M$   $(\mathbf{y}_n)_{n \geq 1}$  reçu par le réseau de capteurs peut se mettre sous la forme

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n,$$

où  $\mathbf{A} = [\mathbf{a}_M(\theta_1), \dots, \mathbf{a}_M(\theta_K)]$  est la matrice  $M \times K$  collectant les  $K$  vecteurs directionnels de dimension  $M$   $\mathbf{a}_M(\theta_1), \dots, \mathbf{a}_M(\theta_K)$ , où  $\theta_1, \dots, \theta_K$  sont les angles d'arrivée des sources et où  $\mathbf{a}_M(\theta) = \frac{1}{\sqrt{M}}[1, \dots, e^{i(M-1)\theta}]^T$ . Les composantes du vecteur  $\mathbf{s}_n \in \mathbb{C}^K$  coïncident avec les signaux sources reçus à l'instant  $n$ , supposés déterministes et non observables.  $(\mathbf{v}_n)_{n \geq 1}$  est un bruit blanc gaussien complexe dont la matrice de covariance est  $\mathbb{E}[\mathbf{v}_n \mathbf{v}_n^*] = \sigma^2 \mathbf{I}$ . Nous supposons que les vecteurs  $\mathbf{y}_n$  sont observés entre les instants 1 et  $N$ , et nous nous intéressons spécifiquement à des contextes dans lesquels le nombre d'observations  $N$  est nettement plus petit que le nombre de capteurs  $M$ . Dans ce contexte, les classiques méthodes de type sous-espace ne fonctionnent plus, et il est d'usage d'utiliser une approche, qualifiée de lissage spatial, consistant à augmenter artificiellement le nombre d'observations (voir par exemple [14], [11]), et qui, à l'origine, avait été introduite pour améliorer les méthodes sous-espaces dans le case de signaux sources totalement corrélés. Si  $L < M$ , le lissage spatial consiste à former  $L$  sous réseaux d'antennes de  $M - L + 1$  capteurs se recouvrant. A chaque instant  $n$ , l'observation disponible sur le réseau  $l$  est le vecteur  $\mathbf{y}_n^{(l)} = (\mathbf{y}_{l,n}, \dots, \mathbf{y}_{M-L+1,n})^T$ , de sorte que l'on fait apparaître  $NL$  observations de dimension  $M - L + 1$ , c'est-à-dire les vecteurs  $(\mathbf{y}_n^{(l)})_{l=1, \dots, L, n=1, \dots, N}$ . Ces observations sont collectées dans la matrice Hankel par bloc de dimensions  $(M - L + 1) \times NL$   $\mathbf{Y}_N^{(L)}$  définie par

$$\mathbf{Y}_N^{(L)} = (\mathbf{y}_1^{(1)}, \dots, \mathbf{y}_1^{(L)}, \dots, \mathbf{y}_N^{(1)}, \dots, \mathbf{y}_N^{(L)})$$

Afin de bien comprendre la façon dont les signaux sources s'organisent au sein de  $\mathbf{Y}_N^{(L)}$ , nous définissons la matrice de Hankel de rang 1  $\mathcal{A}_L(\theta)$  donnée par

$$\mathcal{A}_L^{(L)}(\theta) = \sqrt{L(M - L + 1)/M} \mathbf{a}_{M-L+1}(\theta) (\mathbf{a}_L(\theta))^T \quad (4.5)$$

et la matrice  $(M - L + 1) \times KL$  de rang  $K$   $\mathbf{A}^{(L)}$  s'écrivant

$$\mathbf{A}^{(L)} = \left( \mathcal{A}^{(L)}(\theta_1), \mathcal{A}^{(L)}(\theta_2), \dots, \mathcal{A}^{(L)}(\theta_K) \right) \quad (4.6)$$

$\mathbf{Y}_N^{(L)}$  peut alors se mettre sous la forme

$$\mathbf{Y}_N^{(L)} = \mathbf{A}^{(L)} (\mathbf{S}_N \otimes \mathbf{I}_L) + \mathbf{V}_N^{(L)} \quad (4.7)$$

où  $\mathbf{S}_N$  est la matrice  $K \times N$   $\mathbf{S}_N = (\mathbf{s}_1, \dots, \mathbf{s}_N)$  supposée de rang  $K$ , et où  $\otimes$  représente le produit de Kronecker. Il est facile de vérifier que  $\mathbf{A}^{(L)} (\mathbf{S}_N \otimes \mathbf{I}_L)$  est de rang  $K$ , et que son espace image coïncide avec l'espace de dimension  $K$  engendré par les vecteurs directionnels de dimension  $M - L + 1$   $\mathbf{a}_{M-L+1}(\theta_1), \dots, \mathbf{a}_{M-L+1}(\theta_K)$ . Quand  $NL$  est suffisamment grand et que  $M - L + 1$  est suffisamment petit, la loi des grands nombres permet de se convaincre que la matrice de covariance empirique  $\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*} / NL$  se comporte comme  $\mathbf{A}^{(L)} (\mathbf{S}\mathbf{S}^* \otimes \mathbf{I}_L / NL) \mathbf{A}^{(L)*} + \sigma^2 \mathbf{I}_{M-L+1}$ . On peut alors utiliser l'approche sous-espace pour estimer les  $(\theta_k)_{k=1, \dots, K}$ . Plus précisément, désignons par  $\mathbf{\Pi}^{(L)}$  la matrice de projection orthogonale sur ce que l'on appelle l'espace bruit, c'est-à-dire le complément orthogonal de  $\text{sp}\{\mathbf{a}_{M-L+1}(\theta_1), \dots, \mathbf{a}_{M-L+1}(\theta_K)\}$ , qui coïncide aussi avec l'espace propre associé à la plus petite valeur propre (égale à  $\sigma^2$ ) de  $\mathbf{A}^{(L)} (\mathbf{S}\mathbf{S}^* \otimes \mathbf{I}_L / NL) \mathbf{A}^{(L)*} + \sigma^2 \mathbf{I}_{M-L+1}$ . Soit  $\eta(\theta)$  le pseudo-spectre défini par

$$\eta(\theta) = \mathbf{a}_{M-L+1}(\theta)^* \mathbf{\Pi}^{(L)} \mathbf{a}_{M-L+1}(\theta)$$

Alors, les angles  $(\theta_k)_{k=1, \dots, K}$  sont les solutions uniques de l'équation  $\eta(\theta) = 0$ . En pratique, la méthode sous-espace traditionnelle consiste à remplacer  $\mathbf{\Pi}^{(L)}$  par la matrice de projection orthogonale  $\hat{\mathbf{\Pi}}_N^{(L)}$  sur l'espace propre associé aux  $M - L + 1 - K$  plus petites valeurs propres de  $\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*} / NL$ , et à estimer les angles comme les  $K$  minimas les plus significatifs de la fonction  $\theta \rightarrow \hat{\eta}_N^{(t)}(\theta)$  définie par

$$\hat{\eta}_N^{(t)}(\theta) = \mathbf{a}_{M-L+1}(\theta)^* \hat{\mathbf{\Pi}}_N^{(L)} \mathbf{a}_{M-L+1}(\theta)$$

Quand  $NL \rightarrow +\infty$  et que  $M - L + 1$  reste fixe, la loi des grands nombres implique que  $\|\hat{\mathbf{\Pi}}_N^{(L)} - \mathbf{\Pi}^{(L)}\| \rightarrow 0$  et que

$$\sup_{\theta \in [-\pi, \pi]} \left| \hat{\eta}_N^{(t)}(\theta) - \eta(\theta) \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0. \quad (4.8)$$

Cela implique alors que les estimateurs  $(\hat{\theta}_{k,N}^{(t)})_{k=1, \dots, K}$  définis par

$$\hat{\theta}_{k,N}^{(t)} = \text{argument} \min_{\theta \in \mathcal{I}_k} |\hat{\eta}_N^{(t)}(\theta)|,$$

où les  $(\mathcal{I}_k)_{k=1, \dots, K}$  sont des intervalles disjoints contenant les  $(\theta_k)_{k=1, \dots, K}$ , sont consistants, c'est-à-dire que

$$\hat{\theta}_{k,N}^{(t)} \xrightarrow[N \rightarrow \infty]{a.s.} \theta_k. \quad (4.9)$$

Quand le nombre d'antennes  $M$  est grand, le régime asymptotique dans lequel  $NL \rightarrow +\infty$  et  $M - L + 1$  fixe n'est pas très intéressant en pratique car il implique évidemment que l'antenne virtuelle associée à l'approche de lissage spatial est composée d'un nombre de capteurs nettement plus petit que  $M$ , ce qui limite considérablement la résolution de la méthode sous-espace correspondante. Nous étudions donc la méthode sous-espace dans le cas où les dimensions  $M - L + 1$  et  $NL$  de la matrice  $\mathbf{Y}_N^{(L)}$  sont du même ordre de grandeur et que  $L \ll M$  afin que le schéma de lissage spatial ne conduise pas à une réduction significative du nombre de capteurs. Afin de formaliser cela, nous considérons le régime asymptotique suivant. Nous supposons que  $N$  et  $L$  dépendent de  $M$  et que

$$M \rightarrow +\infty, N = \mathcal{O}(M^\beta), 1/3 < \beta \leq 1, \frac{M - L + 1}{NL} \rightarrow c \quad (4.10)$$

où  $0 < c < \infty$ . Dans la suite, le symbole  $N \rightarrow +\infty$  sera employé pour représenter le régime (4.10). Dans la suite, nous allons montrer que les résultats de [8] permettent de mettre en évidence des estimateurs de type sous-espace consistants

des angles d'arrivée. Puisque  $\frac{M-L+1}{NL} \rightarrow c$ , il est clair que  $\frac{L}{M} \rightarrow 0$  et que  $L = \mathcal{O}(M^\alpha)$  où  $\alpha = 1 - \beta$  vérifie  $0 \leq \alpha < 2/3$ . Par conséquent, le régime (4.10) est similaire à (2.2) si ce n'est qu'il convient d'échanger  $(M, N)$  avec  $(N, M - L + 1)$  pour passer de (2.2) à (4.10). Par ailleurs, si  $\mathbf{W}_N$  est la matrice définie par (3.4), on peut constater que l'échange  $(M, N) \rightarrow (N, M - L + 1)$  fait que les propriétés de la matrice  $\sqrt{\frac{N}{ML}} \mathbf{W}_N^*$  sont identiques à celles de  $\frac{1}{\sqrt{NL}} \mathbf{V}_N^{(L)}$ . Dès lors, les résultats de [8] peuvent être utilisés pour analyser les propriétés des valeurs propres de  $\frac{1}{NL} \mathbf{V}_N^{(L)} \mathbf{V}_N^{(L)*}$ , qui, en gros se comportent comme si les éléments de  $\mathbf{V}_N^{(L)}$  étaient des variables aléatoires i.i.d. En particulier, la distribution empirique des valeurs propres de  $\frac{1}{NL} \mathbf{V}_N^{(L)} \mathbf{V}_N^{(L)*}$  converge vers la distribution de Marcenko-Pastur de paramètres  $(\sigma^2, c)$ , et les valeurs propres restent localisées au voisinage de  $[\sigma^2(1 - \sqrt{c})^2, \sigma^2(1 + \sqrt{c})^2]$ .

Il convient maintenant d'étudier le comportement des  $K$  plus grandes valeurs propres et vecteurs propres associés de la matrice de covariance empirique  $\frac{1}{NL} \mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}$  où d'après (4.7),  $\mathbf{Y}_N^{(L)}$  peut être vue comme une perturbation déterministe de rang  $K$  de la matrice  $\mathbf{V}_N^{(L)}$ . Si  $\mathbf{V}_N^{(L)}$  était à éléments i.i.d., les résultats de [5] permettraient de caractériser le comportement des  $K$  plus grandes valeurs propres et vecteurs propres de  $\frac{1}{NL} \mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}$ . On peut alors établir que les résultats de [8], quelque peu enrichis, impliquent que les  $K$  plus grandes valeurs propres et vecteurs propres de  $\frac{1}{NL} \mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}$  se comportent également comme si  $\mathbf{V}_N^{(L)}$  était à éléments i.i.d. En particulier, on peut établir le résultat suivant.

**Theorem 1** *Supposons que les  $K$  valeurs propres non nulles  $(\lambda_{k,N})_{k=1,\dots,K}$  de  $\frac{1}{NL} \mathbf{A}^{(L)} (\mathbf{S}_N \mathbf{S}_N^* \otimes \mathbf{I}_L) \mathbf{A}^{(L)*}$  convergent vers des quantités  $\lambda_1 > \lambda_2 > \dots > \lambda_K$ . Désignons par ailleurs par  $(\hat{\lambda}_{k,N})_{k=1,\dots,M-L+1}$  et  $(\hat{\mathbf{u}}_{k,N})_{k=1,\dots,M-L+1}$  les valeurs propres (ordonnées par ordre décroissant) et vecteurs propres associés de  $\frac{1}{NL} \mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}$ . Soit  $s$ ,  $0 \leq s \leq K$ , le plus grand entier pour lequel  $\lambda_s > \sigma^2 \sqrt{c}$ . Alors, les  $s$  plus grandes valeurs propres  $(\hat{\lambda}_{k,N})_{k=1,\dots,s}$  vérifient*

$$\hat{\lambda}_{k,N} \xrightarrow[N \rightarrow \infty]{a.s.} \rho_k = \phi(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 c)}{\lambda_k} > x^+.$$

tandis que les  $(\hat{\lambda}_{k,N})_{k=s+1,\dots,K}$  convergent presque sûrement vers  $x^+$ . De plus, pour tous vecteurs déterministes de norme 1  $(\mathbf{a}_N)$  et  $(\mathbf{b}_N)$ , pour  $k = 1, \dots, s$ ,

$$\mathbf{a}_N^* (\hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* - h(\rho_k) \mathbf{u}_{k,N} \mathbf{u}_{k,N}^*) \mathbf{b}_N \rightarrow 0 \text{ a.s.} \quad (4.11)$$

où la fonction  $h(z)$  apparaissant dans (4.11) dépend de la transformée de Stieltjes de la distribution de Marcenko-Pastur  $\mu_{\sigma^2,c}$  et a une expression analytique connue.

Le Théorème 1 permet de montrer que la méthode sous-espace développée dans le cas  $L = 1$  dans [7] continue à fonctionner dans le contexte du régime asymptotique (4.10). Plus précisément:

**Theorem 2** *Si la condition de séparation*

$$\lambda_K > \sigma^2 \sqrt{c} \quad (4.12)$$

est vérifiée, l'estimateur  $\hat{\eta}_N(\theta)$  défini par

$$\hat{\eta}_N(\theta) = (\mathbf{a}_{M-L+1}(\theta))^* \left( \mathbf{I} - \sum_{k=1}^K \frac{1}{h(\hat{\lambda}_{k,N})} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{a}_{M-L+1}(\theta). \quad (4.13)$$

vérifie

$$\sup_{\theta \in [-\pi, \pi]} |\hat{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (4.14)$$

Enfin, les estimateurs des angles  $(\hat{\theta}_{k,N})_{k=1,\dots,K}$  obtenus en minimisant  $\hat{\eta}_N(\theta)$  sont consistants, et ils vérifient

$$M(\hat{\theta}_{k,N} - \theta_k) \rightarrow 0 \text{ a.s.} \quad (4.15)$$

Ce théorème montre que si la condition (4.12) est vérifiée, alors on peut estimer les angles d'arrivées par méthode sous-espace avec une vitesse de convergence meilleure que  $\frac{1}{M}$  dans le cas où le nombre de capteurs  $M$  est grand et où le nombre d'observations  $N$  est très inférieur à  $M$  tout en étant suffisamment grand. En pratique, pour  $M, N, L$  finis, la condition (4.12) signifie que  $\lambda_{K,N}$  doit suffisamment éloigné de  $\sigma^2 \sqrt{c_N}$  où  $c_N$  est défini par  $c_N = \frac{M-L+1}{NL}$ , et cela doit être interprété comme un critère permettant de bien distinguer les valeurs propres dues au signal utile des valeurs propres dues au bruit. Il est évidemment intéressant de rendre la condition (4.12) plus explicite. Pour cela, nous supposons que  $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$  converge une matrice diagonale  $\mathbf{D}$ . Les éléments diagonaux  $(d_k)_{k=1,\dots,K}$  de  $\mathbf{D}$  représentent les puissances des différents signaux source. Dans ce cas, si l'on désigne par  $\mathbf{A}_{M-L+1}$  la matrice  $\mathbf{A}_{M-L+1} = (\mathbf{a}_{M-L+1}(\theta_1), \dots, \mathbf{a}_{M-L+1}(\theta_K))$ , il est facile de voir que la matrice  $\frac{1}{NL} \mathbf{A}^{(L)} (\mathbf{S}_N \mathbf{S}_N^* \otimes \mathbf{I}_L) \mathbf{A}^{(L)*}$  se comporte comme  $\mathbf{A}_{M-L+1} \mathbf{D} \mathbf{A}_{M-L+1}^*$ . Si l'on suppose que les angles  $\theta_k$  restent fixes quand  $M, N, L$  augmentent, la matrice  $\mathbf{A}_{M-L+1}^* \mathbf{A}_{M-L+1} \rightarrow \mathbf{I}_K$ , et les valeurs propres de  $\frac{1}{NL} \mathbf{A}^{(L)} (\mathbf{S}_N \mathbf{S}_N^* \otimes \mathbf{I}_L) \mathbf{A}^{(L)*}$  convergent vers les  $(d_k)_{k=1,\dots,K}$ . Dans ces conditions, la condition de séparation (4.12) signifie que les puissances de toutes les sources dépassent le seuil de "détection"  $\sigma^2 \sqrt{c}$ . Pour des valeurs finies de  $M, N, L$ , ceci veut dire que pour tout  $k$ ,

$$\frac{1}{N} \sum_{n=1}^N |s_{k,n}|^2 > \sigma^2 \sqrt{c_N} \quad (4.16)$$

où  $(s_{k,n})_{n=1,\dots,N}$  est le signal transmis par la source  $k$ . Dans le cas où  $L$  reste fini, ou de façon équivalente, que  $M$  et  $N$  sont du même ordre de grandeur, il est intéressant d'étudier la façon dont le seuil de détection  $\sigma^2 \sqrt{c_N}$  évolue quand  $L$  varie.  $c_N$  est égal à  $\frac{M-L+1}{NL}$  et se comporte donc comme  $\frac{1}{L} \frac{M}{N}$  quand  $M$  et  $N$  sont suffisamment grands par rapport à  $L$ . Dès lors, utiliser une technique de lissage spatial avec  $L$  réseaux d'antennes permet d'abaisser le seuil de détection du facteur  $\frac{1}{\sqrt{L}}$ . Cette analyse n'est cependant justifiée que si  $L$  reste très inférieur à  $M$  et  $N$ . Si  $L$  augmente trop, la diminution de  $M - L + 1$  va impliquer que la matrice  $\mathbf{A}_{M-L+1}^* \mathbf{A}_{M-L+1}$  s'éloigne de l'identité, que la plus petite valeur propre de  $\frac{1}{NL} \mathbf{A}^{(L)} (\mathbf{S}_N \mathbf{S}_N^* \otimes \mathbf{I}_L) \mathbf{A}^{(L)*}$  va diminuer et que la condition (4.16) n'est plus une bonne approximation de la condition (4.12). En d'autres termes, on peut s'attendre à ce que les performances des estimateurs commencent à s'améliorer quand  $L$  augmente, mais qu'à partir d'une certaine valeur de  $L$ , elles se dégradent.

Afin d'illustrer ces résultats théoriques, nous présentons des résultats de simulations numériques. Nous considérons le cas de  $K = 2$  sources dont les angles sont séparés de  $\frac{\pi}{2M}$ , et supposons que  $M = 160$  et  $N = 20$ . La matrice  $2 \times N$  des signaux sources est la réalisation d'une matrice aléatoire dont les éléments sont i.i.d. et de loi gaussienne complexe centrée de variance 1. Les signaux sources sont normalisés de telle façon que leur puissance moyenne soit égale à 1. Ainsi, le rapport signal sur bruit est égal à  $\text{SNR} = 1/\sigma^2$ . Le tableau 1 met en évidence la valeur minimale du SNR pour laquelle la condition de séparation, dans sa version à  $M, N, L$  finie, est vérifiée, i.e.

$$(\sigma^2)^{-1} = \frac{1}{\lambda_{K,N}} \sqrt{(M-L+1)/NL}$$

Ainsi que cela a été expliqué, quand  $L$  augmente,  $\sqrt{(M-L+1)/NL}$  décroît, mais  $\lambda_{K,N}$  tend à diminuer car  $M-L+1$  diminue. Cela explique pourquoi le SNR minimal décroît, puis croît quand  $L$  augmente.

Table 1: Valeur minimale du SNR assurant la condition de séparation.

<b>L</b>	2	4	8	16	32	64	96	128
<b>SNR</b>	33.46	30.30	27.46	25.31	24.70	28.25	36.11	51.52

La figure 4 représente l'erreur quadratique moyenne de l'estimateur  $\hat{\theta}_1$  (appelé G-MUSIC SS) pour  $L = 2, 4, 8, 16$  versus SNR. La borne de Cramer-Rao est également représentée. Comme cela était prévisible, les performances tendent à s'améliorer quand  $L$  augmente jusqu'à  $L = 16$ .

Sur la figure 5,  $L$  est égal à 16, 32, 64, 96, 128. Pour  $L = 32$ , l'erreur quadratique moyenne tend à se dégrader à haut SNR par rapport au cas  $L = 16$ , alors que la perte de performance devient sévère pour les plus grandes valeurs de  $L$ .



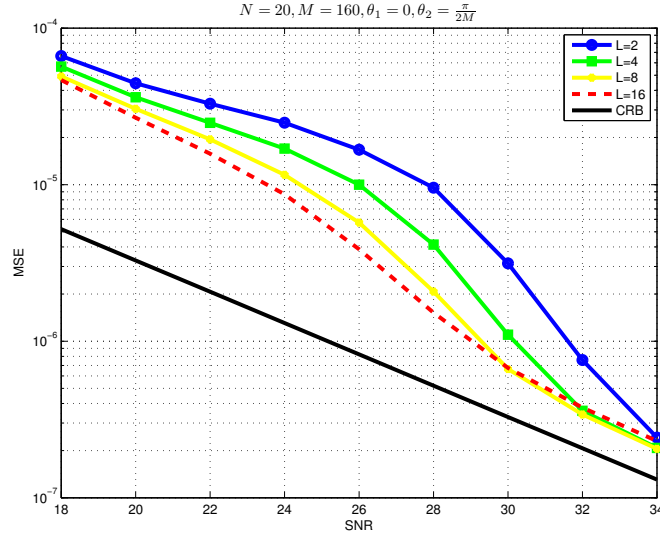


Figure 4: MSE de l'estimateur G-MUSIC SS  $\hat{\theta}_1$  versus SNR

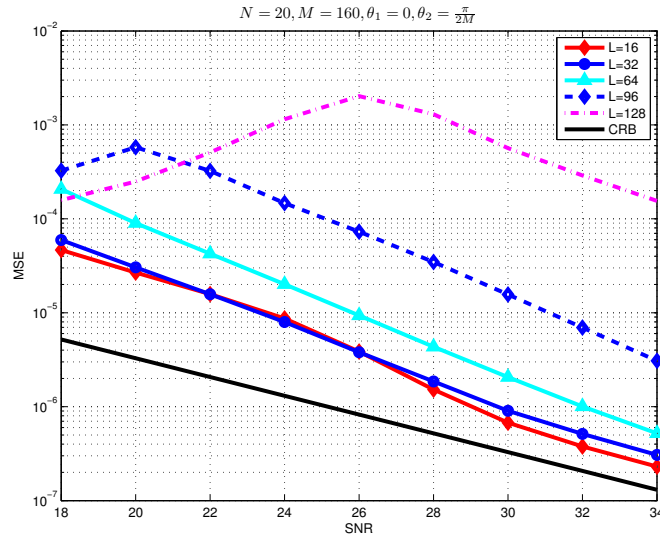


Figure 5: MSE de l'estimateur G-MUSIC SS  $\hat{\theta}_1 \hat{\theta}_1$  versus SNR

Sur la figure 6,  $L$  est fixé à 16. Nous comparons les performances de G-MUSIC SS avec celles de l'algorithme MUSIC standard avec lissage spatial, et représentons également les performances obtenues sans lissage spatial, c'est-à-dire quand  $L = 1$ . Il est clair que l'algorithme G-MUSIC SS permet une très nette amélioration des performances par rapport aux autres.

Nous considérons finalement le cas  $L = 128$ , et comparons comme plus haut G-MUSIC SS, MUSIC SS, G-MUSIC et MUSIC pour  $L = 1$ . G-MUSIC SS a des performances catastrophiques car  $L$  et  $M$  sont du même ordre de grandeur, une situation dans laquelle le Théorème 2 n'est plus valide.

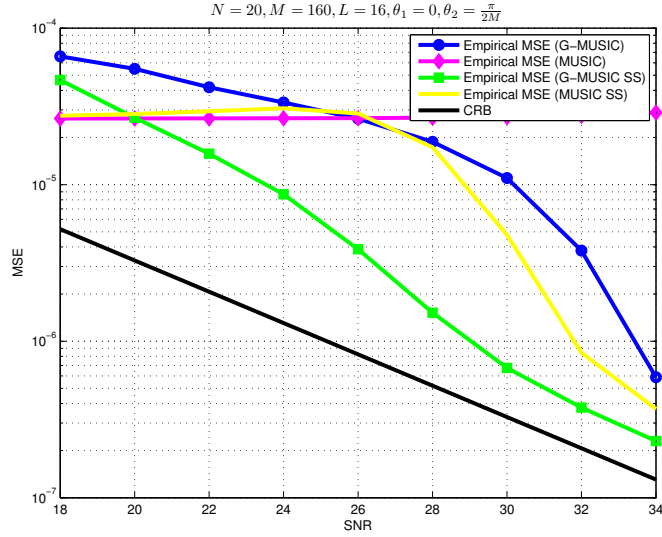


Figure 6: Empirical MSE of different estimators of  $\theta_1$  when  $L=16$

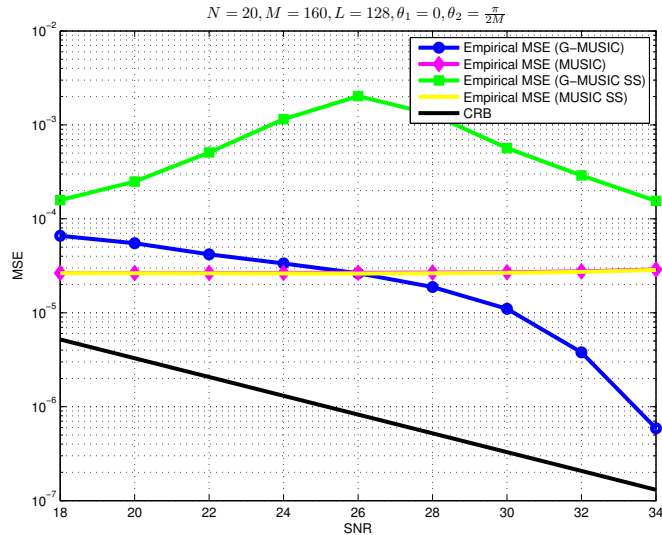


Figure 7: MSE de différents estimateur de  $\theta_1$  when  $L=128$

## 5 Perspectives immédiates.

Nous terminons ce résumé des travaux effectués au titre du WP3 en mettant en évidence quelques perspectives à court terme. Pour cela, nous nous replaçons dans le cadre présenté dans l'introduction, et conservons les notations qui y sont définies. Nous allons nous intéresser à la mise en évidence et à l'étude de statistiques permettant la détection de la source large bande à partir des observation  $(\mathbf{y}_n)_{n=1, \dots, N}$ . Dans ce contexte, le test du rapport du maximum de vraisemblance (GLRT) ne peut être implémenté car les estimateurs au sens du maximum de vraisemblance des vecteurs  $(\mathbf{a}_p)_{p=0, \dots, P-1}$  ne sont pas donnés par une forme explicite. Il convient donc de proposer des tests plus empiriques. En s'inspirant du contexte des sources bande étroite, il est raisonnable de s'intéresser à la statistique de test définie comme la plus grande valeur propre  $\hat{\lambda}_{1,N}$  de la matrice de covariance spatio-temporelle empirique  $\frac{1}{N} \mathbf{Y}_L \mathbf{Y}_L^*$  où  $L$  représente un paramètre qui devra être choisi de façon pertinente. Compte tenu des résultats établis dans [8], l'étude de  $\hat{\lambda}_{1,N}$  dans le régime asymptotique (2.2) va s'avérer simple dans le cas où les paramètres  $P$  et  $L$  ne croissent pas avec  $(M, N)$ . Dans

ce contexte, la matrice  $\mathbf{X}_L$  peut s'écrire sous la forme

$$\mathbf{X}_L = \mathbf{A}_L \mathbf{S}_L$$

où  $\mathbf{A}_L$  est la matrice  $ML \times (P + L)$  définie par

$$\mathbf{A}_L = \begin{pmatrix} \mathbf{a}_{P-1} & \dots & \mathbf{a}_0 & 0 & \dots & 0 \\ 0 & \mathbf{a}_{P-1} & \dots & \mathbf{a}_0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{a}_{P-1} & \dots & \mathbf{a}_0 \end{pmatrix}$$

et où  $\mathbf{S}_L$  est une matrice  $(P+L) \times N$  de Hankel construite à partir des valeurs prises par  $(s_n)_{n \in \mathbb{Z}}$  pour  $n = -(P-1), \dots, N$ . Si  $P + L$  ne croit pas avec  $M$  et  $N$ ,  $\mathbf{X}_L$  est une matrice dont le rang reste fixe et  $\mathbf{Y}_L$  peut donc être interprétée comme une perturbation déterministe de rang fixe de la matrice Hankel par bloc  $\mathbf{V}_L$ . Dans ces conditions, les  $P + L$  plus grandes valeurs propres de  $\frac{1}{N} \mathbf{Y}_L \mathbf{Y}_L^*$  se comportent comme le Théorème 1 le prévoit, c'est-à-dire que leur comportement asymptotique va dépendre de la position des valeurs propres de  $\mathbf{A}_L \frac{\mathbf{S}_L \mathbf{S}_L^*}{N} \mathbf{A}_L^*$  par rapport au seuil de détection  $\sigma^2 \sqrt{c_N}$  où  $c_N = L \frac{M}{N}$ . On peut en déduire assez immédiatement le comportement au premier ordre du test de détection basé sur la comparaison de  $\hat{\lambda}_{1,N}$  à un seuil. Sous l'hypothèse  $H_0$  que le signal est absent,  $\hat{\lambda}_{1,N}$  converge vers  $\sigma^2(1 + \sqrt{c})^2$ . Sous l'hypothèse  $H_1$  que le signal est présent, en admettant que la plus grande valeur propre  $\lambda_{1,N}$  de  $\mathbf{A}_L \frac{\mathbf{S}_L \mathbf{S}_L^*}{N} \mathbf{A}_L^*$  converge vers une quantité  $\lambda_1$ , alors, si  $\lambda_1 \leq \sigma^2 \sqrt{c}$ ,  $\hat{\lambda}_{1,N}$  converge vers  $\sigma^2(1 + \sqrt{c})^2$  comme sous l'hypothèse  $H_0$ , et la détection est impossible. Si par contre  $\lambda_1 > \sigma^2 \sqrt{c}$ ,  $\hat{\lambda}_{1,N}$  converge vers une valeur plus grande que  $\sigma^2(1 + \sqrt{c})^2$ . Ces considérations doivent permettre de comprendre l'influence du choix de  $L$ . Le terme  $c$  est égal à  $Ld$  où  $d = \lim_{N \rightarrow +\infty} \frac{M}{N}$ . Par conséquent, le seuil de détection croît en  $\sqrt{L}$  si  $L$  augmente. Pour que cette augmentation soit bénéfique pour la détection, il faut que  $\lambda_1$ , qui dépend aussi de  $L$ , croisse plus vite que  $\sigma^2 \sqrt{d} \sqrt{L}$ . Cette analyse devrait permettre d'avoir une vision claire des conséquences du choix de  $L$  sur les performances de la détection.

## Références

- [1] K. Abed-Meraim, E. Moulines, Ph. Loubaton, "Prediction error method for second-order blind identification", *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 694-705, March 1997.
- [2] A. Basak, A. Bose, S. Sen, "Limiting spectral distribution of sample autocovariance matrices", to appear in *Bernouilli*, can be downloaded on Arxiv, arXiv:1108.3147v1.
- [3] R. Basu, A. Bose, S. Ganguly, R.S. Hazra, "Limiting spectral distribution of block matrices with Toeplitz block structure", *Statist. and Probab. Lett.*, 82 (2012), no. 7, 1430-1438.
- [4] F. Benaych-Georges, R.R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices", *Adv. in Math.*, vol. 227, no. 1 (2011), 494-521.
- [5] F. Benaych-Georges, R.R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices", *J. Multivariate Anal.*, Vol. 111 (2012), 120-135.
- [6] U. Haagerup and S. Thorbjørnsen. A new application of random matrices:  $\text{Ext}(C_{\text{red}}^*(F_2))$  is not a group. *Ann. of Math. (2)*, 162(2):711-775, 2005.
- [7] Hachem, W. and Loubaton, P. and Mestre, X. and Najim, J. and Vallet, P., "A Subspace Estimator for Fixed Rank Perturbations of Large Random Matrices", *Journal of Multivariate Analysis*, vol. 114, pp. 427-447, 2012.
- [8] P. Loubaton, "On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices", arXiv:1405.2006.

- [9] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.
- [10] E. Moulines, P. Duhamel, J.F. Cardoso, S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters", *IEEE Trans. on Signal Processing*, vol. 43, pp. 516-525, February 1995.
- [11] S.U. Pillai, B.H. Kwon, "Performance analysis of MUSIC type high resolution estimators for direction finding in correlated and coherent scenes", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 8, pp. 1176-1189, 1989.
- [12] H. Schultz, "Non commutative polynomials of independent Gaussian random matrices", *Probab. Theory Relat. Fields* 131, 261-309 (2005)
- [13] A. Sen, B. Virag, "The top eigenvalue of the random Toeplitz matrix and the sine kernel", submitted, can be downloaded on Arxiv, arxiv:1109.5494, 2011.
- [14] T.J. Shan, M. Wax, T. Kailath, "On spatial smoothing for estimation of coherent signals", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 8, pp. 806-811, 1985.
- [15] A.J. Van der Veen, S. Talwar, A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems", *IEEE Trans. on Signal Processing*, vol. 45, no. 1, January 1997.
- [16] A.J. Van der Veen, M. Vanderveen, A. Paulraj, "Joint angle and delay estimation using shift-invariant techniques", *IEEE Trans. on Signal Processing*, vol. 46, no. 2, pp. 405-418, February 1998.