

Projet ANR-12-MONU-0003 DIONISOS.

Rapport D2.1-1:

Detection et estimation pour les sources bande étroite.

1 Introduction.

Le présent rapport résume les travaux effectués dans le cadre du WP2 consacré aux problèmes de détection et estimation pour les sources bande étroite. Nos travaux ont concerné l'étude de méthodes sous-espace d'estimation des angles d'arrivées de sources bande étroite utilisant un grand réseau de capteurs, la détection et l'estimation en bruit corrélé, les tests de sphéricité en grande dimension, l'analyse d'estimateurs robustes de matrice de covariance appelés M -estimateurs, et l'étude des propriétés de l'estimateur standard du tenseur des cumulants d'ordre 4 en grande dimension qui est lié au comportement d'algorithmes de séparation aveugle de sources. Ces travaux ont donné lieu aux publications ou aux pré-publications suivantes:

Articles de revue

- W. Hachem, A. Hardy and J. Najim, "Large Complex Correlated Wishart Matrices: Fluctuations and Asymptotic Independence at the Edges". Submitted, September 2014.
- J. Vinogradova, R. Couillet, and W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection", submitted to *IEEE Transactions on Signal Processing* in March 2014. Revised in July 2014.
- W. Hachem and R. Couillet, "Analysis of the limiting spectral measure of large random matrices of the separable covariance type", accepted for publication in *Random Matrices: Theory and Applications* in October 2014.
- J. Vinogradova, R. Couillet and W. Hachem, "Statistical Inference in Large Antenna Arrays under Unknown Noise Pattern", *IEEE Transactions on Signal Processing*, Vol. 61, no. 22, pp. 5633-5645, 2013.
- C. Butucea, R. Zgheib, "Sharp minimax tests for large covariance matrices", submitted, available on Arxiv, arxiv:1409.1429.

Articles de conférence.

- J. Vinogradova, R. Couillet and W. Hachem, "A new method for source detection, power estimation, and localization in large sensor networks under noise with unknown statistics", *Proc. IEEE ICASSP* 2013.
- Julia Vinogradova, Romain Couillet, Walid Hachem, "Nouvelle méthode de détection de sources, d'estimation de puissances et de localisation dans un système de communication sans fil avec des statistiques de bruit inconnues", *Actes du colloque GRETSI*, 2013.
- Julia Vinogradova, Romain Couillet, Walid Hachem, "Estimation of large Toeplitz covariance matrices and application to source detection", *Proc. EUSIPCO* 2014, Lisbon, Portugal.
- Romain Couillet, Frédéric Pascal, "Robust M-estimator of scatter for large elliptical samples", *Actes du IEEE Workshop on Statistical Signal Processing*, Gold Coast, Australie, 2014.

- P. Vallet, P. Loubaton, " Toeplitz rectification and DoA estimation with MUSIC, in Proceedings ICASSP 2014, pp. pp. 2237-2241, Florence, May 2014.
- P. Vallet, P. Loubaton, X. Mestre " A statistical comparison between MUSIC and G-MUSIC", soumis à ICASSP 2015.

2 Méthodes sous-espace d'estimation des angles d'arrivées de sources bande étroite utilisant un grand réseau de capteurs.

Dans cette section, nous résumons des travaux permettant de compléter l'analyse des performances statistiques de méthodes classiques de localisation de source bande étroite dans le cas où le nombre de capteurs et le nombre d'observations sont grands et comparables.

2.1 Rappel de la problématique.

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

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n, \quad (2.1)$$

où $\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)]$ est la matrice $M \times K$ collectant les K vecteurs directionnels de dimension M $\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)$, où $\theta_1, \dots, \theta_K$ sont les angles d'arrivée des sources et où $\mathbf{a}(\theta)$ est donné par

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}} [1, \dots, e^{i(M-1)\theta}]^T \quad (2.2)$$

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 au problème de l'estimation des angles d'arrivée $(\theta_k)_{k=1, \dots, K}$ à partir de la donnée de la matrice $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$. Nous supposons dans la suite de cette section que le nombre de capteurs M et le nombre d'observations N sont grands et comparables, ce que l'on formalise classiquement par le régime asymptotique:

$$M \rightarrow +\infty, N \rightarrow +\infty, \frac{M}{N} \rightarrow c \quad (2.3)$$

où le paramètre c est strictement positif. Le nombre de sources K est par contre supposé fixe, ce qui, en pratique, modélise des situations dans lesquelles $K/M \ll 1$. Un certain nombre de travaux ont été consacrés depuis 2008 à l'analyse des méthodes d'estimation de type sous-espace dans le régime (2.3). Afin de présenter brièvement les principes de ce type de méthode, appelons $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_N)$ la matrice collectant les K signaux sources, supposée de rang K . Les méthodes sous-espace utilisent le fait que les contributions des signaux sources aux observations, c'est-à-dire les vecteurs $\mathbf{A}\mathbf{s}_1, \dots, \mathbf{A}\mathbf{s}_N$ vivent dans l'espace de dimension K engendré par les vecteurs directions $(\mathbf{a}(\theta_k))_{k=1, \dots, K}$, habituellement appelé "Espace Signal". Dans ces conditions, les angles d'arrivée $\theta_1, \dots, \theta_K$ sont les uniques zéros de ce que l'on appelle le "pseudo-spectre spatial", c'est-à-dire la fonction $\theta \rightarrow \eta(\theta)$ donnée par

$$\eta(\theta) = \mathbf{a}(\theta)^* \mathbf{\Pi} \mathbf{a}(\theta), \quad (2.4)$$

$\mathbf{\Pi}$ représente la matrice de projection orthogonale sur "l'Espace Bruit", défini comme le complément orthogonal de l'espace signal, ou de façon équivalente, comme le noyau (de dimension $M - K$) de la matrice $N^{-1} \mathbf{A} \mathbf{S} \mathbf{S}^* \mathbf{A}^*$. $\mathbf{\Pi}$ n'étant pas connu, il doit être estimé à partir de la matrice \mathbf{Y} . Dans les méthodes sous-espace standard, on utilise l'estimateur

$\hat{\Pi}$ défini comme le projecteur sur l'espace propre associé aux $M - K$ plus petites valeurs propres de la matrice de covariance empirique

$$\frac{\mathbf{Y}\mathbf{Y}^*}{N} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^*,$$

La méthode "MUSIC" consiste alors à estimer les angles $\theta_1, \dots, \theta_K$ comme les K minimas locaux les plus significatifs de l'estimateur traditionnel du pseudo-spectre spatial défini par

$$\hat{\eta}_N^{(t)}(\theta) = \mathbf{a}(\theta)^* \hat{\Pi} \mathbf{a}(\theta), \quad (2.5)$$

où le symbole $^{(t)}$ signifie traditionnel. Cette démarche est justifiée dans le cas où $N \rightarrow +\infty$ tandis que M est fixe car, dans ce régime asymptotique, la loi des grands nombres implique que

$$\left\| \frac{\mathbf{Y}\mathbf{Y}^*}{N} - \left(\mathbf{A} \frac{\mathbf{S}\mathbf{S}^*}{N} \mathbf{A}^* + \sigma^2 \mathbf{I} \right) \right\| \rightarrow 0, \quad (2.6)$$

où $\|\cdot\|$ représente la norme spectrale. Ceci implique que

$$\left\| \hat{\Pi} - \Pi \right\| \xrightarrow[N \rightarrow \infty]{a.s.} 0 \quad (2.7)$$

et que

$$\sup_{\theta \in [-\pi, \pi]} \left| \hat{\eta}_N^{(t)}(\theta) - \eta(\theta) \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0. \quad (2.8)$$

Les K estimateurs fournis par la méthode MUSIC, formellement définis par

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

où les \mathcal{I}_k sont des intervalles disjoints contenant les θ_k sont alors consistants, en ce sens que

$$\hat{\theta}_{k,N} \xrightarrow[N \rightarrow \infty]{a.s.} \theta_k.$$

Dans le cadre du régime asymptotique (2.3), (2.6) et par conséquent, (2.7) et (2.8), ne sont pas vérifiés. La pertinence de la méthode sous-espace en grandes dimensions est donc tout à fait discutable. C'est la raison pour laquelle des travaux visant à proposer de nouvelles méthodes sous-espace dans le contexte des grandes dimensions ont été proposées depuis 2008 (voir le premier article [8], et certains autres qui ont suivis [12], [6]). La contribution [6] utilise un résultat de [3] s'intéressant au comportement des plus grandes valeurs singulières et vecteur singuliers associés d'une grande matrice aléatoire définie comme la somme d'une matrice à éléments indépendants identiquement distribués (i.i.d) avec une matrice déterministe de petit rang. Dans le contexte qui est le nôtre, le résultat [3] est appliqué à la matrice \mathbf{Y} qui peut se mettre sous la forme

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{V} \quad (2.9)$$

où $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_N)$ est une matrice aléatoire à éléments i.i.d., et où $\mathbf{A}\mathbf{S}$ est de rang K qui est supposé rester fixe quand M et N croissent.

Théorème 1 *Supposons que les K valeurs propres non nulles $(\lambda_{k,N})_{k=1,\dots,K}$ de $\mathbf{A} \frac{\mathbf{S}\mathbf{S}^*}{N} \mathbf{A}^*$ convergent vers des quantités $\lambda_1 > \lambda_2 > \dots > \lambda_K$. Désignons par ailleurs par $(\hat{\lambda}_{k,N})_{k=1,\dots,M}$ et $(\hat{\mathbf{u}}_{k,N})_{k=1,\dots,M}$ les valeurs propres (ordonnées par ordre décroissant) et vecteurs propres associés de la matrice de covariance empirique $\frac{1}{N} \mathbf{Y}\mathbf{Y}^*$. 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*

$$\hat{\lambda}_{k,N} \xrightarrow[N \rightarrow \infty]{a.s.} \rho_k = \phi(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 c)}{\lambda_k} > \sigma^2 (1 + \sqrt{c})^2.$$

tandis que les $(\hat{\lambda}_{k,N})_{k=s+1,\dots,K}$ convergent presque sûrement vers $\sigma^2(1+\sqrt{c})^2$. De plus, pour tous vecteurs déterministes de norme 1 (\mathbf{a}_N) et (\mathbf{b}_N) , pour $k = 1, \dots, s$,

$$\mathbf{a}_N^* \left(\hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* - h(\rho_k) \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{b}_N \rightarrow 0 \text{ a.s.} \quad (2.10)$$

où la fonction $h(z)$ apparaissant dans (2.10) est une fonction dont la forme analytique est connue.

La fonction $h(z)$ dépend de la transformée de Stieltjes de la distribution de Marcenko-Pastur $\mu_{\sigma^2,c}$ qui joue un rôle clé dans toutes ces questions. $\mu_{\sigma^2,c}$ a été introduite dans [7], et représente une loi de probabilité qui possède une densité portée par l'intervalle $[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$ (et 0 si $c > 1$). Dans le régime (2.3), la distribution empirique des valeurs propres de la matrice $\frac{1}{N} \mathbf{V} \mathbf{V}^*$ converge vers $\mu_{\sigma^2,c}$, ce qui, intuitivement, signifie que l'histogramme des valeurs propres de toute réalisation de $\frac{1}{N} \mathbf{V} \mathbf{V}^*$ s'accumule autour du graphe de la densité de $\mu_{\sigma^2,c}$. Par ailleurs, il s'avère que si M et N sont assez grands, alors toutes les valeurs propres non nulles de $\frac{1}{N} \mathbf{V} \mathbf{V}^*$ restent localisées au voisinage de $[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$. Ce phénomène, très important, permet de comprendre de façon quantitative le fait intuitif que si le rapport M/N n'est pas très petit, les valeurs propres de la matrice $\frac{1}{N} \mathbf{V} \mathbf{V}^*$, qui tend vers $\sigma^2 \mathbf{I}$ dans le régime asymptotique standard M fixe et $N \rightarrow +\infty$, ont une tendance à s'étaler autour de σ^2 . Le théorème 1 signifie que certaines des valeurs propres de $\frac{1}{N} \mathbf{Y} \mathbf{Y}^*$ peuvent dépasser la plus grande valeur propre $\sigma^2(1+\sqrt{c})^2$ due au bruit si les valeurs propres correspondantes de $\mathbf{A} \frac{\mathbf{S} \mathbf{S}^*}{N} \mathbf{A}^*$ sont plus grandes que le seuil $\sigma^2 \sqrt{c}$. Nous illustrons ce comportement par la figure 1 qui représente l'histogramme des valeurs propres de $\frac{1}{N} \mathbf{Y} \mathbf{Y}^*$ dans le cas où $K = 2$. Les 2 valeurs propres de $\mathbf{A} \frac{\mathbf{S} \mathbf{S}^*}{N} \mathbf{A}^*$ sont suffisamment grandes, de sorte que 2 valeurs propres de la matrice de covariance empirique sortent du support de la distribution de Marcenko-Pastur.

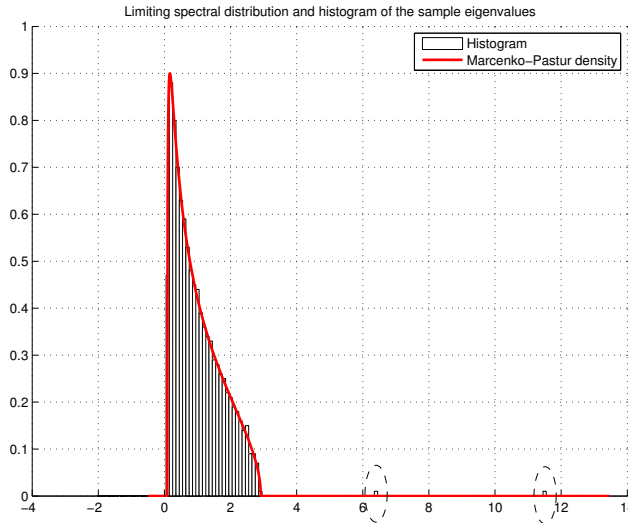


Figure 1: Valeurs propres de $\frac{1}{N} \mathbf{Y} \mathbf{Y}^*$

L'article [6] profite de ce résultat pour établir le théorème suivant.

Théorème 2 Si la condition de séparation

$$\lambda_K > \sigma^2 \sqrt{c} \quad (2.11)$$

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

$$\hat{\eta}(\theta) = \mathbf{a}(\theta)^* \left(\mathbf{I} - \sum_{k=1}^K \frac{1}{h(\hat{\lambda}_{k,N})} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{a}(\theta). \quad (2.12)$$

vérifie

$$\sup_{\theta \in [-\pi, \pi]} |\hat{\eta}_N(\theta) - \eta(\theta)| \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (2.13)$$

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

$$M(\hat{\theta}_{k,N} - \theta_k) \rightarrow 0 \text{ a.s.} \quad (2.14)$$

et

$$N^{3/2}(\hat{\theta}_{k,N} - \theta_k) \rightarrow_{\mathcal{D}} \mathcal{N}(0, \kappa_k) \quad (2.15)$$

pour une certaine valeur de κ , où $\mathcal{N}(0, \kappa_k)$ représente une variable aléatoire gaussienne centrée de variance κ_k , et où le symbole \mathcal{D} signifie la convergence en loi.

Ce théorème montre que si la condition (2.11) est vérifiée, alors, il est possible d'estimer de façon consistante, au sens de la norme de la convergence uniforme, le pseudo-spectre $\theta \rightarrow \eta(\theta)$, ainsi que les angles d'arrivées. Les estimateurs $(\hat{\theta}_{k,N})_{k=1,\dots,K}$ sont appelés estimateurs G-MUSIC.

2.2 Progrès dans l'analyse des méthodes sous-espaces obtenus dans le cadre du WP2.

Le Théorème 2 est établi dans [6] en supposant que les K angles $(\theta_k)_{k=1,\dots,K}$ restent fixes quand M et N augmentent. En pratique, ceci modélise un contexte dans lequel les angles sont éloignés par rapport à la limite de résolution $\frac{2\pi}{M}$ du réseau d'antennes. Dans l'article [14], disponible sur le site du projet, nous montrons que dans ce cadre, sous réserve que la condition (2.11) soit vérifiée, alors la méthode MUSIC traditionnelle fournit également des estimateurs consistants des angles, et qu'ils vérifient (2.15) avec la même variance κ_k . En d'autres termes, bien que l'estimateur traditionnel du pseudo-spectre spatial $\eta_N^{(t)}(\theta)$ défini par (2.5) ne soit pas consistant, les angles qui le minimisent sont des estimateurs consistants qui ont les mêmes performances que celles des estimateurs G-MUSIC. Ceci est illustré par la figure 2 qui représente l'erreur quadratique moyenne des estimateurs MUSIC et G-MUSIC du premier angle θ_1 dans le cas où $K = 2$, $M = 40$, $N = 80$, et où $\theta_2 = \theta_1 + 5 \times \frac{2\pi}{M}$. La borne de Cramer-Rao est également représentée. Dans cet exemple, la différence entre les 2 angles est relativement conséquente, et on peut constater que les estimateurs MUSIC et G-MUSIC ont des performances très proches.

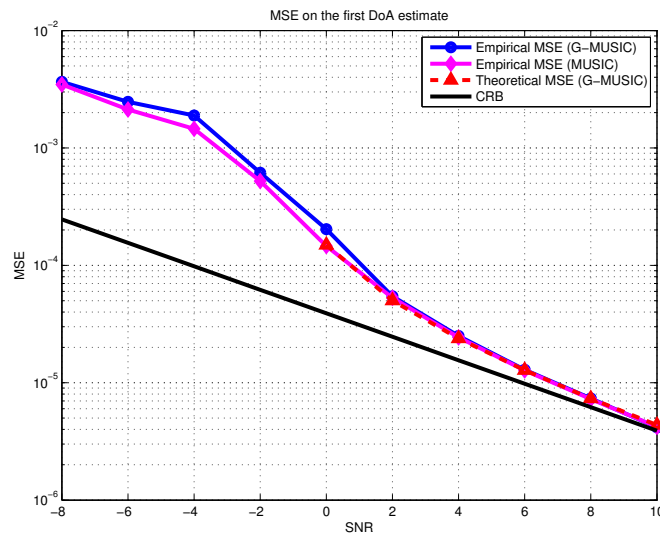


Figure 2: MSE des estimateurs $\hat{\theta}_{1,N}$ and $\hat{\theta}_{1,N}^{(t)}$ pour des angles espacés

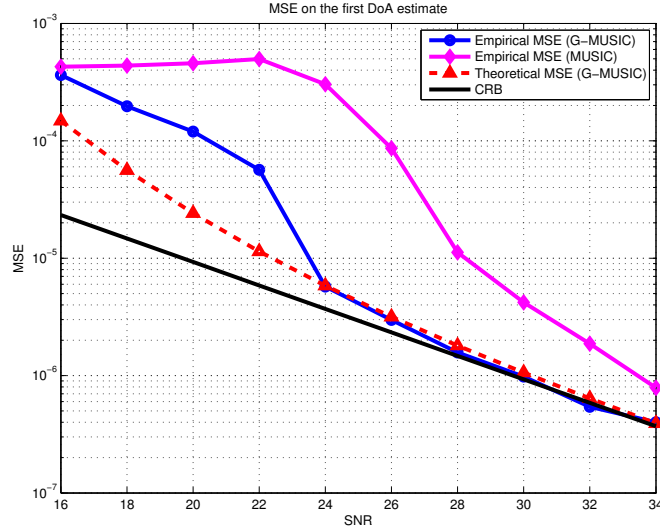


Figure 3: MSE des estimateurs $\hat{\theta}_{1,N}$ and $\hat{\theta}_{1,N}^{(t)}$ pour des angles proches

Le cas d'angles éloignés n'est évidemment pas toujours le plus intéressant puisque l'un des intérêts des méthodes de type MUSIC est d'être à haute résolution, c'est-à-dire qu'elles ont la capacité à estimer dans de bonnes conditions 2 angles qui diffèrent d'un terme $\mathcal{O}(\frac{1}{M})$. L'article [14] s'intéresse également au cas où les angles se rapprochent les uns des autres à la vitesse $\frac{1}{M}$. Dans [14], nous considérons le cas $K = 2$ et $\frac{1}{N}\mathbf{S}\mathbf{S}^* \rightarrow \mathbf{I}_2$, et nous supposons que $\theta_2 = \theta_1 + \alpha/M$. Dans ce contexte, nous montrons que si la condition (2.11), équivalente à

$$\left| \text{sinc}\left(\frac{\alpha}{2}\right) \right| < 1 - \sigma^2 \sqrt{c},$$

est vérifiée, alors les estimateurs G-MUSIC vérifient (2.14) et (2.15) (avec des variances κ_k qui diffèrent de celles du cas d'angles lointains). Par contre, les estimateurs MUSIC traditionnels perdent en général la propriété (2.14) et exhibent donc une vitesse de convergence plus faible. Ceci est attesté par la figure 3 dans laquelle $\theta_2 = \theta_1 + \frac{\pi}{2M}$, de sorte que les 2 angles sont 20 fois plus rapprochés que dans le contexte de la figure 2.

2.3 Influence de la Toeplitzification de la matrice de covariance empirique sur les performances des méthodes sous-espace.

Dans le contexte du modèle (2.1), dans le cas où les signaux sources vérifient

$$\frac{\mathbf{S}\mathbf{S}^*}{N} \xrightarrow{N \rightarrow \infty} \mathbf{\Gamma}$$

où $\mathbf{\Gamma} = \text{Diag}(\gamma_1, \dots, \gamma_K)$ est une matrice diagonale, la matrice $\mathbf{A} \frac{\mathbf{S}\mathbf{S}^*}{N} \mathbf{A}^*$ a le même comportement que

$$\sum_{k=1}^K \gamma_k \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^*$$

qui est une matrice de Toeplitz du fait de la structure (2.2). Dans ces conditions, si M reste fixe et que N tend vers l'infini, (2.6) implique que la matrice de covariance empirique $\frac{\mathbf{Y}\mathbf{Y}^*}{N}$ est proche d'une matrice de Toeplitz. Il est donc pertinent de remplacer $\frac{\mathbf{Y}\mathbf{Y}^*}{N}$ par la matrice de Toeplitz $\hat{\mathbf{R}}$ obtenue en remplaçant chaque terme de $\frac{\mathbf{Y}\mathbf{Y}^*}{N}$ par la moyenne des éléments situés sur la diagonale à laquelle il appartient. En d'autres termes, l'élément (i, j) de $\hat{\mathbf{R}}$ est donné par

$$[\hat{\mathbf{R}}]_{i,j} = \frac{1}{M - |i - j|} \sum_{k-l=i-j} \left[\frac{\mathbf{Y}\mathbf{Y}^*}{N} \right]_{k,l}.$$

Nous nous sommes intéressés dans [13] au comportement de la matrice $\hat{\mathbf{R}}$, et aux propriétés des estimateurs de type sous-espace qui sont obtenus en remplaçant la matrice de covariance empirique par $\hat{\mathbf{R}}$. Ces estimateurs sont appelés dans la suite estimateurs R-MUSIC. Le principal résultat de [13] est que dans le régime asymptotique (2.3),

$$\left\| \hat{\mathbf{R}} - \left(\mathbf{A} \frac{\mathbf{S}\mathbf{S}^*}{N} \mathbf{A}^* + \sigma^2 \mathbf{I} \right) \right\| \rightarrow 0, \quad (2.16)$$

Dans ces conditions, les estimateurs R-MUSIC obtenus à partir de $\hat{\mathbf{R}}$ sont consistants quelque soit le rapport signal sur bruit. Leurs performances asymptotiques sont étudiées dans [13] où il est établi que leur variance décroissent à la vitesse $\frac{1}{N^3}$ comme les estimateurs G-MUSIC. Le principal intérêt des estimateurs R-MUSIC est que la condition de séparation (2.11) n'est plus nécessaire pour assurer leur consistance. Par contre, les performances sont essentiellement limitées par la vitesse avec laquelle la matrice $\frac{\mathbf{S}\mathbf{S}^*}{N}$ converge vers une matrice diagonale. En pratique, les estimateurs basés sur $\hat{\mathbf{R}}$ se comportent mieux que les estimateurs G-MUSIC quand le rapport signal sur bruit (SNR) est faible, mais leurs performances saturent quand le SNR augmente du fait que $\frac{\mathbf{S}\mathbf{S}^*}{N}$ n'est pas exactement une matrice de Toeplitz. La figure 4 illustre ce comportement dans le cas $K = 2$, $M = 20$, $N = 40$, et où $\theta_1 - \theta_2 = \frac{\pi}{2M}$

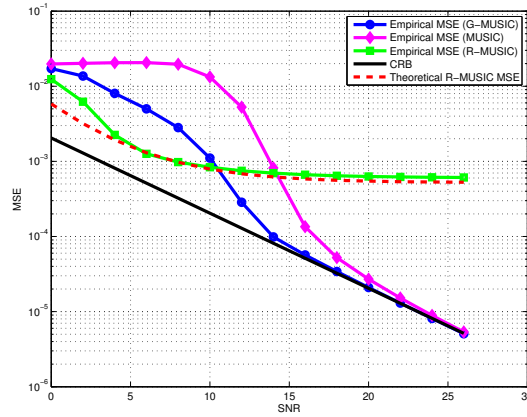


Figure 4: MSE des estimateurs MUSIC, G-MUSIC, R-MUSIC.

3 Détection et estimation en bruit corrélé.

Nous revenons au modèle (2.1) et nous supposons maintenant que le bruit soit *corrélé*. La matrice du bruit \mathbf{V} (cf. Eq. (2.9)) s'écrit maintenant $\mathbf{V} = \mathbf{W}\mathbf{R}^{1/2}$ où \mathbf{W} est une matrice $M \times N$ à éléments gaussiens complexes circulaires standard indépendants, \mathbf{R} est une matrice de covariance $N \times N$ et $\mathbf{R}^{1/2}$ est une "racine carrée" quelconque de \mathbf{R} . Ce modèle est celui d'un bruit blanc spatialement et corrélé temporellement (le bruit corrélé spatialement et blanc temporellement se traite de la même manière). Si la matrice \mathbf{R} était connue du récepteur, il suffirait de remplacer la matrice du signal reçu \mathbf{Y} par la matrice $\mathbf{Y}\mathbf{R}^{-1/2}$ – on parle de "blanchiment" du signal reçu – pour revenir à la situation du paragraphe précédent. On s'intéresse ici au cas où cette matrice n'est pas connue par le récepteur. Nous considérerons les problèmes suivants :

1. Détection du nombre de sources, de leurs puissances et de leurs directions d'arrivées dans le cas où la dimension K du sous-espace signal est fixe,
2. Toujours dans le cas où K est fixe, on s'intéresse à la situation où le bruit est stationnaire, *i.e.*, la matrice \mathbf{R} est de Toeplitz. Nous nous intéressons dans ce cadre aux performances des algorithmes basés sur la Toeplitzification de la matrice de covariance empirique $N^{-1}\mathbf{Y}\mathbf{Y}^*$. Remarquons que la Toeplitzification de $N^{-1}\mathbf{Y}\mathbf{Y}^*$ est liée ici au caractère Toeplitz de \mathbf{R} , et non pas à celui de la composante signal comme au paragraphe 2.3 de ce document,

3. Cas où K et N sont du même ordre de grandeur. Ici, on sait que si c est suffisamment petit, le support de la mesure spectrale limite de $N^{-1}\mathbf{Y}\mathbf{Y}^*$ est en général constitué de plusieurs composantes connexes. Il est alors possible d'envisager un algorithme d'estimation de K basé sur la détermination du nombre de valeurs propres de $N^{-1}\mathbf{Y}\mathbf{Y}^*$ dans chacune de ces composantes connexes. Afin d'évaluer les performances d'un tel algorithme, il convient d'étudier les *fluctuations asymptotiques* des valeurs propres aux bords de ces composantes,
4. Enfin, la situation plus générale de bruits bi-corrélés *bi corrélés* mérite l'attention. La matrice d'un bruit de ce type est de la forme $\mathbf{V} = \mathbf{R}_s^{1/2}\mathbf{W}\mathbf{R}_t^{1/2}$ où la matrice \mathbf{R}_s (resp. \mathbf{R}_t) est une matrice de covariance spatiale (resp. temporelle). Revenons au cas mono corrélé où l'une des matrices \mathbf{R}_s ou \mathbf{R}_t est une matrice identité. Considérons l'algorithme de G-estimation d'une fonctionnelle des valeurs propres de la matrice de covariance ou l'algorithme G-MUSIC d'estimation d'un sous-espace propre associé à cette matrice. Comme nous l'avons sommairement montré dans le premier rapport d'avancement du projet DIONISOS, la conception de ces algorithmes est intimement liée à la détermination du support de la mesure spectrale limite de $N^{-1}\mathbf{V}\mathbf{V}^*$. Dans cet ordre d'idées, la détermination de ce support dans le cas bi corrélé est un prélude à la conception de G-estimateurs adaptés à ce contexte.

3.1 Le cas où \mathbf{R} est générale et K est fixe

Revenons au modèle décrit par l'équation (2.9) et précisons nos hypothèses :

1. La matrice du bruit s'écrit $\mathbf{V} = \mathbf{W}\mathbf{R}^{1/2}$ où \mathbf{W} est une matrice $M \times N$ à éléments gaussiens complexes circulaires standard indépendants, et où $\mathbf{R} \in \mathbb{C}^{N \times N}$ est une matrice hermitienne semi définie positive dont la mesure spectrale converge faiblement vers une mesure de probabilité limite ν . Le support de ν est l'intervalle $[a_\nu, b_\nu] \subset [0, \infty)$, et $\nu(\{0\}) = 0$. Par ailleurs, en notant par $\sigma_1^2, \dots, \sigma_N^2$ les valeurs propres de \mathbf{R} ,

$$\max_{n \in \{1, \dots, N\}} \text{dist}(\sigma_n^2, \text{support}(\nu)) \xrightarrow{N \rightarrow \infty} 0, \quad (3.17)$$

2. Si $\nu(\{b_\nu\}) = 0$, alors il existe $\varepsilon > 0$ et une fonction $f_\nu(t) \geq C(b_\nu - t)$ sur $[b_\nu - \varepsilon, b_\nu]$, avec $C > 0$, telle que pour tout borélien A de $[a_\nu, b_\nu]$,

$$\nu(A \cap [b_\nu - \varepsilon, b_\nu]) = \int_{A \cap [b_\nu - \varepsilon, b_\nu]} f_\nu(t) dt. \quad (3.18)$$

Cette hypothèse stipule que soit ν met une masse en $\{b_\nu\}$, soit elle possède une densité qui décroît suffisamment vite près de b_ν .

3. Dans le modèle $\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{V}$, la matrice de signal $\mathbf{B} = \mathbf{A}\mathbf{S}$ est aléatoire et indépendante de \mathbf{W} . Par ailleurs, son rang K est fixe et $\sup_N \|\mathbf{B}\| < \infty$,
4. Il existe une factorisation $\mathbf{B} = \mathbf{U}\mathbf{X}$ telle que $\mathbf{U} \in \mathbb{C}^{M \times K}$ satisfait $\mathbf{U}^H\mathbf{U} = \mathbf{I}_K$, et pour $z \in \mathbb{C} - \text{support}(\nu)$,

$$\mathbf{X}(\mathbf{R} - z\mathbf{I}_N)^{-1}\mathbf{X}^H \xrightarrow[N \rightarrow \infty]{\text{a.s.}} m_\nu(z)\mathbf{\Gamma} \quad (3.19)$$

où $m_\nu(z)$ est la transformée de Stieltjes de ν , et $\mathbf{\Gamma} = \text{diag}(\gamma_1\mathbf{I}_{j_1}, \dots, \gamma_t\mathbf{I}_{j_t})$, $\gamma_1 > \dots > \gamma_t > 0$ et $j_1 + \dots + j_t = K$.

Les hypothèses 1 et 2 sont satisfaites par la grande majorité des modèles de bruit rencontrés dans la pratique. Par exemple, si le bruit est temporellement stationnaire, de fonction de covariance sommable et de densité spectrale de puissance $G(\exp(2i\pi f))$, alors la loi ν est déterminée par l'équation

$$\int g(t)\nu(dt) = \int_0^1 g(G(\exp(2i\pi f)))df$$

pour toute fonction continue g , et on peut démontrer sans difficulté que les conditions (3.17) et (3.18) sont satisfaites.

Considérons maintenant l'hypothèse 4, et supposons que le signal soit décrit par le modèle à bande étroite

$$\mathbf{B} = \mathbf{A}\mathbf{S}$$

où $\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)]$ est une matrice de vecteurs directionnels et où $\mathbf{S} = [s_{k,n}]_{k,n=1}^{K,N}$ et la matrice des signaux émis représentés par les variables aléatoires $s_{k,n}$ indépendantes, centrées et qui satisfont

$$\mathbb{E} \left[s_{k_1,n} s_{k_2,n}^* \right]_{k_1,k_2=1}^K = N^{-1/2} \begin{bmatrix} \gamma_1 \mathbf{I}_{j_1} & & \\ & \ddots & \\ & & \gamma_t \mathbf{I}_{j_t} \end{bmatrix}$$

pour tout $n \in \{1, \dots, N\}$. En écrivant $\mathbf{U} = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1/2}$ et $\mathbf{X} = (\mathbf{A}^H \mathbf{A})^{1/2} \mathbf{S}$, on peut également établir la convergence (3.19) en inspectant la structure de \mathbf{A} et en appliquant une loi des grands nombres. Dans ce modèle, le signal est indépendant temporellement et spatialement. Nous signalons sans plus de commentaires que notre approche peut aussi s'adapter au cas où le signal est corrélé temporellement tout en ayant la même fonction de covariance temporelle que celle du bruit. Ce cas survient quand la corrélation temporelle est due à la présence du filtre de réception.

Rappelons pour commencer les résultats concernant le comportement asymptotique des valeurs propres $\lambda_1, \dots, \lambda_N$ de la matrice de covariance $N^{-1} \mathbf{V} \mathbf{V}^H$ associée au bruit seul. Nous savons depuis [7] que la mesure spectrale de cette matrice converge faiblement presque sûrement vers une mesure limite μ , et que la transformée de Stieltjes $m(z)$ de cette dernière est définie pour tout $z \in \mathbb{C}_+ = \{z : \Im z > 0\}$ comme étant l'unique solution dans $m \in \mathbb{C}_+$ de l'équation

$$m = \left(-z + \int \frac{t}{1 + cmt} \nu(dt) \right)^{-1}.$$

Cette mesure de probabilité est de la forme $\mu(dt) = \max(0, 1 - c^{-1}) \delta_0 + f(t)dt$ où $f(t)$ est une densité continue sur $(0, \infty)$. Le support de $f(t)dt$ est par ailleurs un intervalle compact $[a, b] \subset [0, \infty)$ [11]. Enfin, pour tout intervalle $[x_1, x_2] \subset (0, a) \cup (b, \infty)$,

$$\#\{i : \lambda_i \in [x_1, x_2]\} = 0 \quad \text{p.s. pour tous } N \text{ larges} \quad [2]$$

Notre but est de déterminer le nombre de sources K à partir de l'observation des valeurs propres $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_N$ de la matrice de covariance $N^{-1} \mathbf{Y} \mathbf{Y}^H$. Comme la matrice de signal est une perturbation d'ordre fini de la matrice de bruit \mathbf{V} , la mesure spectrale de $N^{-1} \mathbf{Y} \mathbf{Y}^H$ converge vers μ . Mais, tout comme dans le cas où le bruit est blanc, la matrice $N^{-1} \mathbf{Y} \mathbf{Y}^H$ peut avoir quelques valeurs propres à droite de l'intervalle $[a, b]$:

Théorème 3 ([16]) *La fonction $g(x) = m(x)(cxm(x) + c - 1)$ est positive et décroît de $g(b^+)$ vers zéro sur (b, ∞) . Si $\gamma_1 g(b^+) \leq 1$, alors $\hat{\lambda}_1 \rightarrow b$ presque sûrement. Sinon, soit $s \in \{1, \dots, t\}$ le plus grand indice pour lequel $\gamma_s g(b^+) > 1$. Pour $k = 1, \dots, s$, soit ρ_k la solution unique x dans (b, ∞) de l'équation $\gamma_k g(x) = 1$. Alors, pour $i = 1, \dots, s$*

$$\begin{aligned} \hat{\lambda}_{j_1 + \dots + j_{i-1} + 1}, \dots, \hat{\lambda}_{j_1 + \dots + j_i} &\xrightarrow[N \rightarrow \infty]{\text{p.s.}} \rho_i \\ \hat{\lambda}_{j_1 + \dots + j_s + 1} &\xrightarrow[N \rightarrow \infty]{\text{p.s.}} b \end{aligned}$$

en posant $j_0 = 0$. La condition $\gamma_k g(b^+) > 1$ est équivalente à la condition

$$\gamma_k > \left(\int \frac{-m_b}{1 + cm_b t} \nu(dt) \right)^{-1} \quad (3.20)$$

où m_b est l'unique solution dans $(-(cb_\nu)^{-1}, 0)$ de l'équation en m

$$\int \left(\frac{mt}{1 + cmt} \right)^2 \nu(dt) = \frac{1}{c}.$$

Grâce à ce résultat, nous pouvons estimer d'une manière consistante la dimension K du sous-espace signal, pour peu que les puissances des signaux soient suffisantes :

Théorème 4 Soit $s \geq 0$ le plus grand entier pour lequel l'inégalité (3.20) soit satisfaite. Soit $0 < \varepsilon < (\rho_s/b) - 1$ où $\rho_0 = \infty$. Etant donné $L \geq K$, soit

$$\hat{K}_N = \arg \max_{k \in \{0, \dots, L\}} \frac{\hat{\lambda}_k}{\hat{\lambda}_{k+1}} > 1 + \varepsilon$$

où $\hat{\lambda}_0 = \infty$. Alors, pour tous N larges

$$\hat{K}_N = j_1 + \dots + j_s \quad p.s.$$

où $j_0 = 0$.

Dans [16], un algorithme d'estimation des γ_k et des directions d'arrivée θ_k est également proposé. Une analyse au second ordre (fluctuations) est également réalisée.

3.2 Le cas où \mathbf{R} est Toeplitz et K est fixe

Nous supposons toujours que $\mathbf{Y} = \mathbf{B} + \mathbf{WR}^{1/2}$, mais en nous limitant cette fois au cas où \mathbf{R} est une matrice de Toeplitz. Une technique de détection basée sur des outils ne faisant pas partie de la théorie des grandes matrices aléatoires a été développée dans [17].

Rappelons que dans le cas où \mathbf{R} est connue, il est possible de blanchir le signal reçu par $\mathbf{R}^{-1/2}$ et d'appliquer au signal blanchi le test GLRT ("Generalized Log-likelihood Ratio Test"), qui se ramène dans ce cas à la comparaison de $\|\mathbf{Y}\|^2 / \text{trace}(\mathbf{Y}\mathbf{Y}^H)$ à un seuil.

Supposons maintenant que \mathbf{R} soit inconnue, mais qu'un estimateur $\hat{\mathbf{R}}$ de cette matrice consistant en la norme spectrale soit disponible. Dans ce cas, le test ci-dessus appliqué à $\mathbf{Y}\hat{\mathbf{R}}^{-1/2}$ restera consistant. Dans [17], nous avons montré que quand \mathbf{R} est Toeplitz, de tels estimateurs peuvent être pris dans la famille des estimateurs Toeplitzifiés classiques. Ces derniers s'avèrent en effet consistants en norme spectrale, que le signal de rang fini \mathbf{B} soit présent (à ce stade il constituerait une nuisance) ou non. Afin de mieux motiver ce résultat, nous avons établi des inégalités de concentration sur la norme spectrale de l'erreur d'estimation.

Afin de mieux fixer les idées, considérons dans un premier temps la matrice de bruit pur $\mathbf{V} = [V_{m,n}]_{m,n=1}^{M,N} = \mathbf{WR}^{1/2}$ où

$$\mathbf{R} \triangleq \begin{bmatrix} r_0 & r_1 & \dots & r_{N-1} \\ r_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1 \\ r_{1-N} & \dots & r_{-1} & r_0 \end{bmatrix}.$$

Les estimateurs classiques des coefficients de \mathbf{R} , pour $k = -(N-1), \dots, N-1$, sont donnés par les équations

$$\hat{r}_k^b = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N V_{m,n+k} V_{m,n}^* \mathbb{1}_{0 \leq n+k \leq N}$$

$$\hat{r}_k^u = \frac{1}{M(N-|k|)} \sum_{m=1}^M \sum_{n=1}^N V_{m,n+k} V_{m,n}^* \mathbb{1}_{1 \leq n+k \leq N}.$$

L'estimateur \hat{r}_k^b est biaisé et l'estimateur \hat{r}_k^u ne l'est pas. Les matrices

$$\hat{\mathbf{R}}^b = (\hat{r}_{i-j}^b)_{0 \leq i, j \leq N-1}$$

$$\hat{\mathbf{R}}^u = (\hat{r}_{i-j}^u)_{0 \leq i, j \leq N-1}.$$

construites à partir de ces estimateurs sont structurellement Toeplitz. Ces estimées sont bien connues. Cependant, des inégalités de concentration sur la norme spectrale de l'erreur n'ont pas encore été proposées.

Nous avons montré que pour tout $x > 0$, et dans le régime où M et N tendent vers l'infini à la même vitesse,

$$\mathbb{P} \left[\|\widehat{\mathbf{R}}^b - \mathbf{R}\| > x \right] \leq \exp(-cNf(x)(1+o(1))) \quad (3.21)$$

$$\mathbb{P} \left[\|\widehat{\mathbf{R}}^u - \mathbf{R}\| > x \right] \leq \exp\left(-\frac{cNg(x)}{\log N}(1+o(1))\right) \quad (3.22)$$

où $f(x)$ et $g(x)$ sont des fonctions croissantes et indépendantes de M et de N . Ces inégalités nous montrent que les erreurs d'estimation décroissent exponentiellement quand $N \rightarrow \infty$ et qu'en particulier, les estimées $\widehat{\mathbf{R}}^b$ et $\widehat{\mathbf{R}}^u$ sont consistantes en norme spectrale.

Nous revenons maintenant au modèle "signal plus bruit" $\mathbf{Y} = \mathbf{B} + \mathbf{W}\mathbf{R}^{1/2}$ et nous supposons que le rang de \mathbf{B} est $K = 1$. Les coefficients de \mathbf{R} sont maintenant estimés directement à partir de \mathbf{Y} . Nous avons démontré dans [17] que les inégalités (3.21) et (3.22) restent vraies avec les mêmes fonctions de taux, grâce au fait que le rang de \mathbf{B} est fixe. Grâce à ces résultats, nous pouvons appliquer le détecteur à seuil du GLRT au signal

$$\mathbf{Y}\widehat{\mathbf{R}}^{-1/2} = \mathbf{B}\widehat{\mathbf{R}}^{-1/2} + \mathbf{V}\mathbf{R}^{1/2}\widehat{\mathbf{R}}^{-1/2} \quad (3.23)$$

où $\widehat{\mathbf{R}} = \widehat{\mathbf{R}}^b$ ou $\widehat{\mathbf{R}}^u$. Cette méthode et celle de [16] s'avèrent complémentaires dans le sens où la méthode basée sur la Toeplitzification présente de meilleures performances dans le régime du faible rapport signal à bruit.

3.3 Matrice de covariance générale et $K = \mathcal{O}(N)$: fluctuations et indépendance asymptotiques aux bords du spectre

Dans le modèle $\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{R}^{1/2}\mathbf{W}$, nous supposons maintenant que \mathbf{A} soit déterministe et que la matrice $[\mathbf{S}^T \ \mathbf{W}^T]$ de taille $N \times (K + M)$ soit à éléments gaussiens standard indépendants. Grâce à l'invariance unitaire de la loi gaussienne standard multidimensionnelle, il n'est pas difficile de montrer que $\mathbf{Y}\mathbf{Y}^H$ est égale en loi à la matrice $\boldsymbol{\Sigma}^{1/2}\mathbf{Z}\mathbf{Z}^H(\boldsymbol{\Sigma}^{1/2})^H$, où $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H + \mathbf{R}$ et où \mathbf{Z} est à éléments gaussiens standard. Supposons un instant que \mathbf{R} soit de la forme $\mathbf{R} = \sigma^2\mathbf{I}_M$. La matrice de covariance d'ensemble $\boldsymbol{\Sigma}$ possédera alors $M - K$ valeurs propres égales à σ^2 et K valeurs propres plus grandes que σ^2 . Si c est suffisamment petit, le spectre de la matrice de covariance empirique $N^{-1}\mathbf{Y}\mathbf{Y}^H$ sera le plus souvent composé de deux "groupes" disjoints de valeurs propres. Mathématiquement, ces groupes de valeurs propres correspondent aux composantes connexes de la mesure spectrale limite de $N^{-1}\mathbf{Y}\mathbf{Y}^H$ (ou, si cette mesure limite n'existe pas, de l'équivalent déterministe à M, N finis de la mesure spectrale de $N^{-1}\mathbf{Y}\mathbf{Y}^H$). La plupart des algorithmes d'estimation de K sont basés sur cette observation : en désignant par $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_M$ les valeurs propres de $N^{-1}\mathbf{Y}\mathbf{Y}^H$, bon nombre de ces algorithmes sont basés sur la comparaison de $\hat{\lambda}_k/\hat{\lambda}_{k+1}$ à un seuil. Le rapport le plus élevé entre ces valeurs propres successives est observé lors du passage d'un groupe de valeurs propres à un autre.

Une manière d'évaluer les performances statistiques de ces tests consiste alors à étudier les fluctuations asymptotiques des valeurs propres aux bords des composantes connexes de la mesure spectrale limite.

D'une manière plus générale, considérons le modèle matriciel $\mathbf{Y}\mathbf{Y}^H/N = \boldsymbol{\Sigma}^{1/2}(\mathbf{Z}\mathbf{Z}^H/N)(\boldsymbol{\Sigma}^{1/2})^H$ où $\boldsymbol{\Sigma}$ possède une mesure spectrale limite dont le support peut consister en plusieurs composantes connexes. Le support de la mesure limite μ de $N^{-1}\mathbf{Y}\mathbf{Y}^H$ peut alors consister également en plusieurs composantes connexes ; ce support a été entièrement déterminé dans [11]. Les problèmes que nous nous posons sont alors les suivants. Dans le cas où les éléments de la matrice \mathbf{Y} sont gaussiens,

1. Caractériser les fluctuations asymptotiques des valeurs propres de $N^{-1}\mathbf{Y}\mathbf{Y}^H$ aux bords de toutes les composantes connexes de μ ,
2. Etablir l'indépendance asymptotique de ces valeurs propres extrêmes,
3. Caractériser les fluctuations asymptotiques de la plus petite valeur propre de $N^{-1}\mathbf{Y}\mathbf{Y}^H$ dans le cas où $c = 1$.

Le premier problème a été abordé par Baik *et.al.* et par El Karoui dans [18, 19]. Ces auteurs ont montré que sous une condition de régularité, la plus grande valeur propre de $N^{-1}\mathbf{Y}\mathbf{Y}^H$ converge en loi à l'échelle $N^{2/3}$ vers la loi dite de Tracy-Widom. Dans [20], nous avons généralisé ces résultats aux bords droit et gauche positifs de toutes les composantes connexes du support de μ .

Sous la même condition de régularité, nous avons également établi l'indépendance asymptotique de ces valeurs propres extrêmes. Pour établir ce résultat, nous avons adopté l'approche par opérateurs de Bornemann [21].

Enfin, on sait que la plus petite valeur propre non triviale de $N^{-1}\mathbf{Y}\mathbf{Y}^H$ converge vers zéro uniquement quand $c = 1$. Dans cette situation, le régime des fluctuations de cette valeur propre s'avère différent du régime "Tracy-Widom". Dans [20], nous avons montré que ces fluctuations sont décrites par le noyau de Bessel.

Ce travail nous a permis en particulier d'établir les fluctuations asymptotiques du *nombre de conditionnement* de la matrice $N^{-1}\mathbf{Y}\mathbf{Y}^H$, notamment dans le cas où $c = 1$.

3.4 Le cas bicorrélé : détermination du support

Considérons le modèle $\mathbf{H} = N^{-1}\mathbf{R}_s^{1/2}\mathbf{W}\mathbf{R}_t\mathbf{W}^H(\mathbf{R}_t^{1/2})^H$ où \mathbf{R}_s et \mathbf{R}_t sont des matrices déterministes semi définies positives et où \mathbf{W} de dimensions $M \times N$ est à éléments indépendants et identiquement distribués. Quand \mathbf{R}_s et \mathbf{R}_t possèdent des mesures spectrales limite ν et $\tilde{\nu}$ et quand $M/N \rightarrow c > 0$ quand $N \rightarrow \infty$, la mesure spectrale de \mathbf{H} converge faiblement au sens presque sûr vers une mesure limite μ dont la transformée de Stieltjes est définie comme étant la solution unique d'un système d'équations (voir entre autres [23, 15]). Dans l'article [22], nous avons étudié les propriétés de μ . Ce travail reprend en les généralisant toutes les conclusions de l'article [11] qui traite du cas mono corrélé. Les principales conclusions de [22] sont les suivantes :

- La mesure μ possède une densité sur $(0, \infty)$. Cette densité est analytique partout où elle est positive,
- Le support de μ peut être déterminé aisément à partir de ν , $\tilde{\nu}$ et c ,
- En général, la densité de μ se comporte en $\sqrt{|x - a|}$ près d'un bord a d'une composante connexe du support de μ . La constante associée à cette décroissance en $\sqrt{|x - a|}$ est calculée. On conjecture que ce comportement en $\sqrt{|x - a|}$ est lié à des fluctuations asymptotiques des valeurs propres extrêmes de \mathbf{H} de type Tracy-Widom.

4 Test de sphéricité à rayon connu en grande dimension

Nous résumons dans ce paragraphe les résultats obtenus dans [24] concernant l'analyse de tests de sphéricité en grande dimension. La problématique des tests de sphéricité est de tester de façon non paramétrique que des observations Gaussiennes i.i.d. de dimension M $\mathbf{y}_1, \dots, \mathbf{y}_N$ ont une matrice de covariance $\mathbf{\Sigma}$ qui est un multiple de l'identité. Cela permet d'aborder le problème de la détection de sources bande étroite dans un bruit gaussien blanc temporellement et spatialement car l'hypothèse absence de signal correspond au cas où $\mathbf{\Sigma}$ est un multiple de l'identité. Dans le cas où M et N tendent vers l'infini au même rythme, la problématique de la détection de sources bande étroite a souvent été abordée en faisant l'hypothèse que le nombre de sources est très petit devant M et N (voir par exemple [25] ainsi que [16], [17] réalisés au cours du projet). L'un des avantages de la formulation liée aux tests de sphéricité est de ne nécessiter aucune hypothèse sur le nombre de sources.

Dans le travail [24], nous supposons, comme dans [25], que les signaux sources sont des suites i.i.d. gaussiennes indépendantes entre elles, et que les éléments de la matrice de covariance spatiale due aux signaux sources décroît vers 0 suffisamment vite lorsque l'on s'éloigne de la diagonale.

Plus formellement, on observe les vecteurs i.i.d. Gaussiens réels de dimension M $\mathbf{y}_1, \dots, \mathbf{y}_N$, et l'on souhaite tester

que $\Sigma = \mathbf{I}_M$ contre l'hypothèse que $\Sigma = (\sigma_{ij})_{(i,j)=1,\dots,M}$ vérifie

$$\begin{aligned} H_0 &: \Sigma = \mathbf{I}_M \\ H_1(\alpha, L, \varphi) &: \Sigma \text{ telle que } \frac{1}{2M} \sum_{i \neq j} \sigma_{ij}^2 \geq \varphi^2 \text{ et } \frac{1}{2M} \sum_{i \neq j} \sigma_{ij}^2 |i - j|^{2\alpha} \leq L, \text{ pour tout } M, \end{aligned}$$

pour $\alpha > 1/2, L, \varphi > 0$. Dans les tests de sphéricité les plus généraux, l'hypothèse H_0 postule que Σ est un multiple, appelé rayon, inconnu de \mathbf{I}_M , mais nous considérons ici le cas où le rayon est connu. Comme ce paramètre représente en pratique le niveau du bruit, il n'est pas absurde de le supposer connu. La quantité $\varphi = \varphi(M, N)$ qui peut dépendre de M et N , contrôle la distance entre les matrices définies par l'hypothèse H_1 et \mathbf{I}_M , et est appelé vitesse de séparation (des hypothèses). L'un des enjeux est de déterminer un paramètre φ optimal en un certain sens. Dans [24], nous abordons le problème via la théorie minimax des tests. Nous nous intéressons à la valeur de séparation φ optimale au sens où on peut trouver une méthode dont le risque de test tend vers 0 uniformément en Σ appartenant à l'alternative (bornes supérieures); par ailleurs, nous démontrons qu'aucune procédure ne peut distinguer Σ et \mathbf{I} quand la distance entre les deux est inférieure à cette valeur optimale (borne inférieure). Nos résultats permettent de décrire cette vitesse de séparation de manière exacte comme étant

$$\tilde{\varphi} = \left(C^{1/2}(\alpha, L) \cdot N\sqrt{M} \right)^{-\frac{2\alpha}{4\alpha+1}}, \quad C(\alpha, L) = \frac{(2\alpha + 1)L^{-1/(2\alpha)}}{2\alpha(4\alpha + 1)^{1+1/(2\alpha)}}.$$

La procédure de test est basée sur une statistique de test D_N qui est une fonctionnelle quadratique pondérée avec des poids w_{ij}^* optimaux qui permettent d'atteindre les bonnes vitesses et constantes associées, voir [24],

$$D_N = \frac{1}{MN(N-1)} \sum_{1 \leq k \neq l \leq N} \sum_{1 \leq i < j \leq M} w_{ij}^* \mathbf{y}_{k,i} \mathbf{y}_{k,j} \mathbf{y}_{l,i} \mathbf{y}_{l,j}.$$

Théorème 5 (*Bornes supérieures*) Si $\varphi \rightarrow 0$ et $M\varphi^{1/\alpha} \rightarrow \infty$ quand M et N tendent vers l'infini, tel que $\liminf \varphi/\tilde{\varphi} > 1$, alors, la statistique de test $\Delta^*(t) = I(D_N > t)$ avec D_N défini précédemment et $t = C^{1/2}(\alpha, L)\varphi^{2+1/(2\alpha)}/2$ vérifie

$$P_I(\Delta^*(t) = 1) + \sup_{\Sigma \in H_1(\alpha, L, \varphi)} P_\Sigma(\Delta^*(t) = 0) \leq 2\Phi\left(-\frac{N\sqrt{M}}{2} C^{1/2}(\alpha, L)\varphi^{2+1/(2\alpha)}\right) + o(1).$$

où Φ représente la fonction de répartition de la loi gaussienne standard. Notons que la condition $M\tilde{\varphi}^{1/\alpha} \rightarrow \infty$ implique une relation entre N et M : $N = o(M^{2\alpha})$. Ceci n'exclut pas des valeurs de M plus grandes que N dès que $\alpha > 1/2$.

Théorème 6 (*Bornes inférieures*) Si $\varphi \rightarrow 0$ quand N et M tendent vers l'infini, tel que $\limsup \varphi/\tilde{\varphi} < 1$, alors

$$\inf_{\Delta} \left\{ P_I(\Delta = 1) + \sup_{\Sigma \in H_1(\alpha, L, \varphi)} P_\Sigma(\Delta = 0) \right\} \geq 2\Phi\left(-\frac{N\sqrt{M}}{2} C^{1/2}(\alpha, L)\varphi^{2+1/(2\alpha)}\right) + o(1),$$

où l'inf est pris sur toutes les fonctions mesurables des observations.

5 M-estimateurs robustes.

5.1 Résultats

L'objectif du projet "M-estimateurs robustes" consiste en la succession de deux étapes comme suit: (i) analyse théorique du comportement asymptotique d'une famille de M-estimateurs de matrices de covariance dits de Maronna lorsque les dimensions du système (nombre d'observations et nombre d'antennes) croissent au même rythme et lorsque les données sont impulsives, et (ii) de cette étude, déduction de méthodes améliorées d'inférence (détection et estimation) robustes à des bruits impulsifs.

Ce premier rapport fait état des avancées théoriques quant à l'étape (i) de cette procédure. Ces progrès ont mené à l'élaboration successive de deux articles. Le premier, qui établit un premier résultat à l'aide d'une preuve non rigoureuse mais plus intuitive, a été publié à la conférence IEEE SSP 2014.

R. Couillet, F. Pascal, "Robust M-estimator of scatter for large elliptical samples", IEEE Workshop on Statistical Signal Processing (SSP'14), Gold Coast, Australia, 2014.

Le second article, qui contient une preuve exhaustive et rigoureuse, a été élaboré dans le cadre d'un projet ERC parallèle et a été publié dans le Elsevier Journal of Multivariate Analysis, comme suit

R. Couillet, F. Pascal, J. W. Silverstein, "The Random Matrix Regime of Maronna's M-estimator with elliptically distributed samples", (à paraître) Elsevier Journal of Multivariate Analysis, 2014.

La question était précisément de considérer le modèle statistique suivante: $x_1, \dots, x_n \in \mathbb{C}^N$ sont des vecteurs définis par $x_i = \sqrt{\tau_i} A_N w_i$, où $\tau_1, \dots, \tau_n \in \mathbb{R}_+$ et $w_1, \dots, w_n \in \mathbb{C}^{\bar{N}}$ sont aléatoires et $A_N \in \mathbb{C}^{N \times \bar{N}}$ est déterministe. On notera $c_N \triangleq N/n$ et $\bar{c}_N \triangleq \bar{N}/N$. Nous demandons précisément que

Assumption 1 Les vecteurs $x_i = \sqrt{\tau_i} A_N w_i$, $i \in \{1, \dots, n\}$, vérifient:

1. la mesure aléatoire $\tilde{\nu}_n = \frac{1}{n} \sum_{i=1}^n \delta_{\tau_i}$ est telle que $\int \tau \tilde{\nu}_n(d\tau) \rightarrow 1$ p.s. et $\int \tau^{1+\varepsilon} \tilde{\nu}_n(d\tau) < \infty$
2. il existe $\varepsilon < 1 - \phi_\infty^{-1} < 1 - c_+$ et $m > 0$ tel que, pour tout n large p.s. $\tilde{\nu}_n([0, m]) < \varepsilon$
3. dénotant $C_N \triangleq A_N A_N^*$, $C_N \succ 0$ et $\limsup_N \|C_N\| < \infty$
4. $w_1, \dots, w_n \in \mathbb{C}^{\bar{N}}$ sont indépendants, unitairement invariant, de moyenne nulle et de norme $\|w_i\|^2 = \bar{N}$; ils sont de plus indépendants de τ_1, \dots, τ_n .

On écrira également $x_i = \tau_i z_i$.

Comme souvent dans ce rapport, nous supposons que

Assumption 2 Pour chaque N , $c_N < 1$, $\bar{c}_N \geq 1$ et

$$c_- < \liminf_n c_N \leq \limsup_n c_N < c_+$$

où $0 < c_- < c_+ < 1$.

On définit alors \hat{C}_N comme l'unique solution de l'équation en Z

$$Z = \frac{1}{n} \sum_{i=1}^n u \left(\frac{1}{N} x_i^* Z^{-1} x_i \right) x_i x_i^* \quad (5.24)$$

où u satisfait les propriétés suivantes:

- (i) $u : [0, \infty) \rightarrow (0, \infty)$ est positive, continue et décroissante
- (ii) $\phi : x \mapsto xu(x)$ est croissante et bornée par $\lim_{x \rightarrow \infty} \phi(x) \triangleq \phi_\infty > 1$
- (iii) $\phi_\infty < c_+^{-1}$.

A l'aide de ces hypothèses, qui miment celles de Maronna dans son article originel, nous obtenons le résultat fondamental suivant:

Théorème 7 (Comportement Asymptotique) Soit \hat{C}_N la solution de (5.24). Alors,

$$\left\| \hat{C}_N - \hat{S}_N \right\| \rightarrow 0$$

p.s. où

$$\hat{S}_N \triangleq \frac{1}{n} \sum_{i=1}^n v(\tau_i \gamma_N) x_i x_i^*$$

et γ_N est l'unique solution de l'équation en γ

$$1 = \frac{1}{n} \sum_{i=1}^n \frac{\psi(\tau_i \gamma)}{1 + c_N \psi(\tau_i \gamma)}$$

avec $v : x \mapsto (u \circ g^{-1})(x)$, $\psi : x \mapsto xv(x)$, et $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $x \mapsto x/(1 - c_N \phi(x))$.

Ce théorème permet de décrire le comportement de la matrice \hat{C}_N , définie uniquement de manière implicite, sous la forme asymptotiquement bien approximée d'une matrice aléatoire \hat{S}_N explicite. Seul γ_N demeure implicite, mais ce dernier n'implique que les variables scalaires τ_i , que l'on peut supposer déterministes.

Etant donné que \hat{S}_N est une matrice aléatoire de modélisation classique et que la convergence en norme spectrale est un résultat très fort, ce théorème permet de caractériser fidèlement le comportement spectral de \hat{C}_N dans le régime $N, n \rightarrow \infty$ pour des données x_i impulsives. Notons en particulier qu'un corollaire de ce résultat est l'assurance de la compacité du support asymptotique de \hat{C}_N , contrairement à $\frac{1}{N} \sum_i x_i x_i^*$ qui sera généralement de support non compact pour des τ_i non bornés. De plus, il est assuré également que la valeur propre maximale de \hat{C}_N reste bornée par $\psi_\infty(1 + \sqrt{c_+})^2/\gamma_+$, où $\gamma_- > 0$ est une borne inférieure uniforme sur γ_N .

Comme la matrice de covariance empirique $\frac{1}{N} \sum_i x_i x_i^*$ et que la matrice \hat{S}_N sous ce modèle des x_i sont deux objets dont il est possible d'extraire des équivalents déterministes pour la loi des valeurs propres, nous obtenons en corollaire la possibilité de décrire théoriquement ces lois approximatives pour les deux modèles, comme suit.

Corollaire 1 (Equivalent Déterministe du Spectre) *Sous les hypothèses du théorème,*

$$\frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(\hat{C}_N)} - \mu_N \rightarrow 0 \quad (5.25)$$

p.s. où la convergence est au sens faible de la convergence en loi, avec μ_N une mesure de probabilité à densité de transformée de Stieltjes $m_N(z)$ donnée, pour $z \in \mathbb{C}_+$, par

$$m_N(z) = -\frac{1}{z} \frac{1}{N} \sum_{i=1}^N \frac{1}{1 + \tilde{\delta}_N(z) \lambda_i(C_N)}$$

où $\tilde{\delta}_N(z)$ est l'unique solution dans \mathbb{C}_+ de l'équation en $\tilde{\delta}$

$$\begin{aligned} \tilde{\delta} &= -\frac{1}{z} \frac{1}{n} \sum_{i=1}^n \frac{\psi(\tau_i \gamma_N)}{\gamma_N + \psi(\tau_i \gamma_N) \tilde{\delta}} \\ \delta &= -\frac{1}{z} \frac{1}{n} \sum_{i=1}^N \frac{\lambda_i(C_N)}{1 + \lambda_i(C_N) \tilde{\delta}} \end{aligned}$$

et γ_N est défini dans l'énoncé du théorème. De plus, le support \mathcal{S}_N de μ_N est borné uniformément. Dans le cas particulier où $C_N = I_N$, $m_N(z)$ est la solution unique pour $z \in \mathbb{C}_+$ de l'équation en m

$$m = \left(-z + \gamma_N^{-1} \frac{1}{n} \sum_{i=1}^n \frac{\psi(\tau_i \gamma_N)}{1 + c \gamma_N^{-1} \psi(\tau_i \gamma_N) m} \right)^{-1}.$$

Par ailleurs, pour chaque $N_0 \in \mathbb{N}$ et chaque fermé $\mathcal{A} \subset \mathbb{R}$ tel que $\mathcal{A} \cap \left(\bigcup_{N \geq N_0} \mathcal{S}_N \right) = \emptyset$,

$$\left| \left\{ \lambda_i(\hat{C}_N) \right\}_{i=1}^N \cap \mathcal{A} \right| \rightarrow 0 \quad (5.26)$$

p.s. de sorte qu'en particulier

$$\limsup_N \|\hat{C}_N\| < \infty. \quad (5.27)$$

Ce corollaire est exploité pour représenter graphiquement les lois des valeurs propres pour $\frac{1}{N} \sum_i x_i x_i^*$ et \hat{C}_N dans les deux figures ci-après pour le cas de variables τ_i i.i.d. de loi à support non compacte. Ces graphes confirment l'observation faite précédemment selon laquelle le support de la loi des valeurs propres de \hat{C}_N est asymptotiquement borné au contraire de celui de la matrice de covariance empirique.

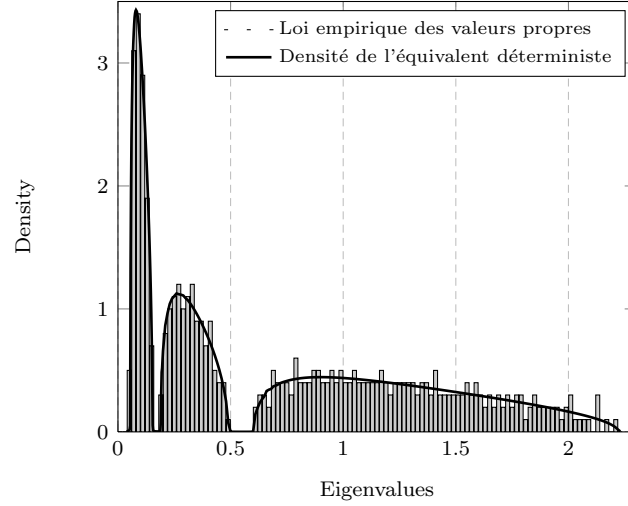


Figure 5: Histogramme des valeurs propres de \hat{C}_N pour $n = 2500$, $N = 500$, $C_N = \text{diag}(I_{125}, 3I_{125}, 10I_{250})$, τ_1 de distribution $\Gamma(.5, 2)$.

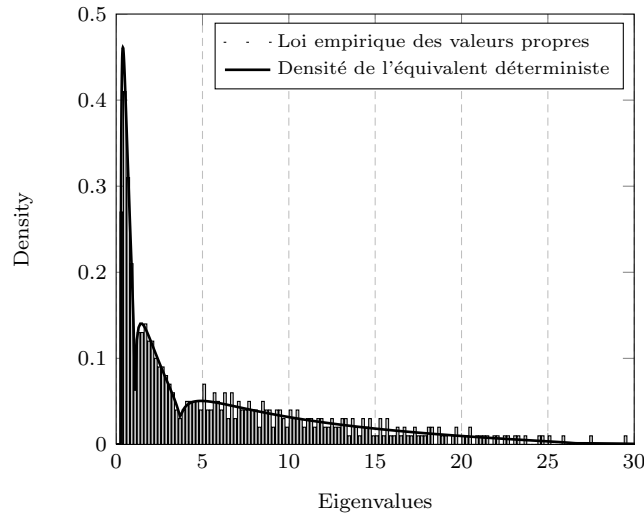


Figure 6: Histogramme des valeurs propres de $\frac{1}{n} \sum_{i=1}^n x_i x_i^*$ pour $n = 2500$, $N = 500$, $C_N = \text{diag}(I_{125}, 3I_{125}, 10I_{250})$, τ_1 de distribution $\Gamma(.5, 2)$.

Dans le cadre d'application pratique où les x_i consistent en des données de bruit (impulsif) pur, aucune valeur propre ne pourra asymptotiquement excéder la limite sus-dite. L'étude que nous menons actuellement consiste en l'analyse du modèle pour lequel un signal déterministe et constant est ajouté au modèle de x_i . Il est attendu que ce signal déterministe, si d'intensité suffisante, génèrera une valeur propre additionnelle dans le spectre de \hat{C}_N que l'on trouvera au-delà de la borne décrite pour le cas bruit seul.

5.2 Idées principales de la preuve

Il convient dans un premier temps de remarquer que, si le résultat est acquis pour $C_N = I_N$, l'extension pour des C_N génériques est triviale. Nous nous plaçons ainsi sous l'hypothèse $C_N = I_N$ à partir d'ici.

Sous cette hypothèse, le résultat du théorème est établi en deux étapes. Une première étape consiste en la réécriture plus intuitive de l'équation définissant \hat{C}_N . Cette réécriture est avantageuse car elle permet d'exhiber des termes (des formes quadratiques) dont une analyse non-rigoureuse permet de maîtriser le comportement. Dans un second temps, cette réécriture étant acquise, une exploitation fine (et cette fois rigoureuse) des propriétés de monotonie des fonctions $u(x)$ et $\phi(x) = xu(x)$ sera menée de manière à obtenir des bornes inférieure et supérieure sur \hat{C}_N dans l'ordre partiel des matrices hermitiennes. En démontrant que les deux bornes, qui s'avèreront être des multiples (respectivement inférieur et supérieur à un) de \hat{S}_N , ont une différence en norme spectrale tendant vers zéro presque sûrement, on obtient le résultat final.

Brièvement, la première étape, purement algébrique permet de démontrer que \hat{C}_N peut être définie de manière équivalente comme

$$\hat{C}_N = \frac{1}{N} \sum_{i=1}^n v(\tau_i d_i) x_i x_i^*$$

où d_1, \dots, d_n forment l'unique solution au système d'équations

$$d_i = \frac{1}{N} z_i^* \hat{C}_{(i)}^{-1} z_i$$

où nous avons défini $\hat{C}_{(i)} = \hat{C}_N - \frac{1}{n} u(\frac{1}{N} x_i^* \hat{C}_N^{-1} x_i) x_i x_i^*$. Nous rappelons que la fonction v est définie dans le corps du théorème. La manipulation algébrique permettant d'obtenir ce résultat consiste simplement en l'observation que la forme quadratique $\frac{1}{N} x_i^* \hat{C}_N^{-1} x_i$ peut s'écrire comme une fonction inversible de $\frac{1}{N} x_i^* \hat{C}_{(i)}^{-1} x_i$ sous les hypothèses posées sur $\phi(x)$.

Le problème se réduit donc ici en l'étude du système de n équations $d_i = \frac{1}{N} z_i^* \hat{C}_{(i)}^{-1} z_i$, $i = 1, \dots, n$. En réécrivant $\hat{C}_{(i)}$ plus simplement sous la forme $\hat{C}_{(i)} = \frac{1}{N} \sum_{j \neq i} \tau_j v(\tau_j d_j) z_j z_j^*$, ce système se réduit donc à l'équation fondamentale:

$$d_i = \frac{1}{N} z_i^* \left(\frac{1}{N} \sum_{j \neq i} \tau_j v(\tau_j d_j) z_j z_j^* \right)^{-1} z_i.$$

A ce niveau, il convient d'introduire la variable $f_i \triangleq v(\tau_i d_i) / v(\tau_i \gamma_N)$ et de réordonner ces f_i (on crée donc ici une structure de dépendence!) de sorte que $f_1 \leq \dots \leq f_n$. L'objectif est de démontrer que les f_i convergent uniformément vers 1. Pour ce faire, on procède par contradiction en supposant premièrement que $f_n > 1 + \ell$ infiniment souvent pour un ℓ positive quelconque. On procédera de même pour f_1 en supposant que $f_1 < 1 - \ell$ infiniment souvent.

Concentrons nous cependant ici sur f_n . Par définition, $v(\tau_j d_j) = f_j v(\tau_j \gamma_N) \leq f_n v(\tau_j \gamma_N)$. En écrivant la définition de $v(\tau_i d_i)$ et en exploitant la décroissance de v (qu'on aura prouvée au préalable), on a alors

$$v(\tau_i d_i) \leq v \left(\tau_i \frac{1}{N} z_i^* \left(\frac{1}{N} \sum_{j \neq i} \tau_j f_n v(\tau_j \gamma_N) z_j z_j^* \right)^{-1} z_i \right)$$

qu'on simplifie alors en

$$v(\tau_i d_i) \leq v \left(\frac{\tau_i}{f_n} \frac{1}{N} z_i^* \left(\frac{1}{N} \sum_{j \neq i} \tau_j v(\tau_j \gamma_N) z_j z_j^* \right)^{-1} z_i \right)$$

Cette manipulation a permis d'isoler la forme quadratique à droite qui est un objet amplement étudié en matrices aléatoires. On montre en particulier ici qu'uniformément sur i , $\frac{1}{N} z_i^* \left(\frac{1}{N} \sum_{j \neq i} \tau_j v(\tau_j \gamma_N) z_j z_j^* \right)^{-1} z_i - \gamma_N \rightarrow 0$ p.s. Ainsi,

en spécialisant $i = n$ au dessus, on a pour des n assez grands,

$$v(\tau_n d_n) \leq v\left(\frac{\tau_n}{f_n}(\gamma_N - \epsilon)\right)$$

où ϵ est fixé mais peut être pris arbitrairement petit (à condition alors d'augmenter n de manière appropriée). En divisant de part et d'autre par $v(\tau_n \gamma_N)$, on a alors

$$f_n \leq \frac{v\left(\frac{\tau_n}{f_n}(\gamma_N - \epsilon)\right)}{v(\tau_n \gamma_N)}.$$

L'exploitation de la croissance de $\phi(x) = xu(x)$ apparait maintenant. En réécrivant l'équation au dessus en fonction de ϕ , on a

$$1 \leq \frac{\phi\left(\frac{\tau_n}{f_n}(\gamma_N - \epsilon)\right)}{\psi(\tau_n \gamma_N)} \frac{\gamma_N}{\gamma_N - \epsilon}.$$

Comme on a fait l'hypothèse que $f_n > 1 + \ell$ infiniment souvent, on peut se réduire encore à une telle suite et obtenir alors, de par la croissance de ϕ ,

$$1 \leq \frac{\phi\left(\frac{\tau_n \gamma_N}{1 + \ell}\right)}{\psi(\tau_n \gamma_N)} \frac{\gamma_N}{\gamma_N - \epsilon}.$$

A ce niveau, s'il est assuré que les τ_i ont un support compact, comme ϵ peut être pris arbitrairement faible, on déduit immédiatement la contradiction $1 < 1$ dans la limite des large n (par exemple en extrayant une sous-séquence convergente de τ_n et en appliquant la limite $n \rightarrow \infty$ sur le membre de droite). Ceci assure ainsi que $f_n \leq 1 + \ell$ pour tout n large p.s. Ayant assuré de manière identique que $f_1 \geq 1 - \ell$ pour tout n large, le théorème se déduit aisément en bornant $v(\tau_i d_i)$ inférieurement et supérieurement par $(1 \pm \ell)v(\tau_i \gamma_N)$ et le résultat s'ensuit.

La réelle difficulté technique consiste à maîtriser le cas (plus intéressant!) où les τ_i ne sont pas de support borné uniformément sur n . En effet, dans ce cas τ_n peut n'avoir qu'un point d'accumulation en $+\infty$ et le membre de droite de l'expression précédente est dans cette limite égal à un, rompant la contradiction. Pour gérer ce cas, la preuve est réinitialisée en prenant $M > 0$ large et en isolant les indices i pour lesquels $\tau_i > M$ et les autres. On opère alors par des arguments de perturbations de petit rang pour montrer qu'on peut restreindre l'analyse précédente aux τ_i de faible amplitude, démontrant ainsi que tous les f_i pour lesquels $\tau_i \leq M$ convergent à nouveau vers un. Les autres f_i ne posent alors aucun problème étant donné qu'ils correspondent à des $\phi(\tau_i d_i)$ qui, en ayant démontré que $\min_i \liminf_n d_i > 0$, tendent vers une constante ϕ_∞ et n'induisent donc aucune perturbation dans l'estimation finale.

6 Analyse d'algorithmes de séparation de sources: étude des propriétés de l'estimateur standard du tenseur des cumulants d'ordre 4 en grande dimension.

6.1 Rappel de la problématique.

Dans le contexte de la séparation aveugle de sources bande étroite, on observe une série temporelle multivariable de dimension M , notée (\mathbf{y}_n) , et qui est supposée être un mélange linéaire de K ($K < M$) signaux aléatoires stationnaires non gaussiens mutuellement indépendants perturbé par un bruit additif gaussien. L'objet de la séparation aveugle de sources est d'estimer la matrice de mélange et les différentes composantes du mélange. Plus formellement, \mathbf{y}_n est défini à chaque instant n par

$$\mathbf{y}_n = \sum_{k=1}^K \mathbf{a}_k s_{k,n} + \mathbf{v}_n = \mathbf{A} \mathbf{s}_n + \mathbf{v}_n$$

où $(s_k)_{k=1,\dots,K}$ sont des suites aléatoires stationnaires non gaussiennes non observables, mutuellement indépendantes, de moyenne nulle et de variance 1, et où les vecteurs de dimension M $(\mathbf{a}_k)_{k=1,\dots,K}$ sont inconnus et supposés déterministes. \mathbf{v} représente un bruit additif gaussien complexe.

\mathbf{y}_n est observé entre les instants $n = 1$ et $n = N$, et le problème est d'identifier la matrice \mathbf{A} de façon à pouvoir, dans un second temps, estimer les signaux non observables $(s_{k,n})_{n=1,\dots,N}$. Pour cela, nous allons nous intéresser au comportement des algorithmes basés sur les cumulants d'ordre 4 du signal observé dans le cas où le nombre de capteurs M est grand. Si on désigne par $(\mathbf{y}_{i,n})_{i=1,\dots,M}$ les composantes de \mathbf{y}_n , alors, l'indépendance mutuelle des sources et la gaussianité du bruit additif implique immédiatement que

$$\text{cum}(\mathbf{y}_{i_1}, \mathbf{y}_{i_2}^*, \mathbf{y}_{i_3}, \mathbf{y}_{i_4}^*) = \sum_{k=1}^K c_4(s_k) \mathbf{a}_{i_1,k} \mathbf{a}_{i_2,k}^* \mathbf{a}_{i_3,k} \mathbf{a}_{i_4,k}^* \quad (6.28)$$

où $c_4(s_k) = \text{cum}(s_k, s_k^*, s_k, s_k^*)$ (nous avons omis de préciser la dépendance temporelle). Si l'on range les divers cumulants dans une matrice $M^2 \times M^2$ \mathbf{C} , (6.28) apparaît équivalent à

$$\mathbf{C} = \sum_{k=1}^K c_4(s_k) (\mathbf{a}_k \otimes \mathbf{a}_k) (\mathbf{a}_k \otimes \mathbf{a}_k)^* \quad (6.29)$$

où \otimes désigne le produit de Kronecker. Par conséquent, la matrice \mathbf{C} est de rang K . Sous certaines hypothèses, (par exemple $c_4(s_k) < 0$ pour $k = 1, \dots, K$), il est possible d'extraire grâce à un algorithme relativement simple les vecteurs $(\mathbf{a}_k)_{k=1,\dots,K}$ à partir des K valeurs propres non nulles et vecteurs propres associés de \mathbf{C} ([1]). En pratique, la matrice \mathbf{C} est estimée par la matrice $\hat{\mathbf{C}}$ définie par

$$\hat{\mathbf{C}} = \frac{1}{N} \sum_{n=1}^N (\mathbf{y}_n \otimes \mathbf{y}_n) (\mathbf{y}_n \otimes \mathbf{y}_n)^* - \frac{1}{N^2} \sum_{(m,n)=1,\dots,N} (\mathbf{y}_m \otimes \mathbf{y}_n) (\mathbf{y}_m \otimes \mathbf{y}_n + \mathbf{y}_n \otimes \mathbf{y}_m)^* \quad (6.30)$$

qui remplace la matrice \mathbf{C} dans l'algorithme précédent. Cette approche fournit des estimateurs consistants des vecteurs $(\mathbf{a}_k)_{k=1,\dots,K}$ dans le cas où M est fixe et que $N \rightarrow +\infty$. L'un des objectifs de la tâche 2-2 est d'étudier le comportement de $\hat{\mathbf{C}}$ dans le cas où M est grand. Ce contexte est modélisé dans la suite par un régime asymptotique dans lequel M et N tendent vers l'infini simultanément. De plus, il convient de préciser que nous supposons dans la suite que lorsque M augmente, les normes des vecteurs $(\mathbf{a}_k)_{k=1,\dots,K}$ restent bornées.

Dans le présent rapport, nous nous intéressons exclusivement aux propriétés de $\hat{\mathbf{C}}$ en absence de signal, c'est-à-dire quand \mathbf{y}_n se réduit au terme de bruit \mathbf{v}_n . Dans ce cas, la vraie matrice des cumulants est $\mathbf{C} = 0$, et notre but est d'évaluer dans quelle mesure $\hat{\mathbf{C}}$ est éloigné de 0 dans les régimes des grandes dimensions. Par ailleurs, une meilleure connaissance des propriétés de $\hat{\mathbf{C}}$ en absence de sources peut être très utile pour mieux appréhender le cas où des signaux sources sont présents. Nous allons supposer que le bruit $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ est blanc temporellement et que $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}$ et $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^T) = 0$. Cette hypothèse permet de simplifier considérablement les évaluations qui suivent, mais il convient de mentionner que les méthodes de séparation de sources auxquelles nous nous intéressons ont pour principal avantage de ne nécessiter aucune hypothèse sur les statistiques du bruit additif. Rien n'empêche toutefois d'étudier des méthodes de séparation de sources générales dans le cas particulier du bruit blanc temporellement et spatialement afin de se faire une idée des problèmes posés par les grandes dimensions. Il est évidemment pertinent de s'intéresser à des modèles de bruit plus généraux, par exemple blancs temporellement, mais pas spatialement.

6.2 Quelques résultats algébriques simples.

Dans la suite, nous allons noter L l'entier

$$L = \frac{M(M+1)}{2} \quad (6.31)$$

Par ailleurs, nous désignons par $\mathbf{\Pi}$ la matrice $M^2 \times M^2$ définie par

$$\mathbf{\Pi}(\mathbf{x}_1 \otimes \mathbf{x}_2) = \mathbf{x}_2 \otimes \mathbf{x}_1 \quad (6.32)$$

pour tous vecteurs $\mathbf{x}_1, \mathbf{x}_2$ de dimensions M . Il est facile de vérifier que $\mathbf{\Pi}$ peut se mettre sous la forme

$$\mathbf{\Pi} = \mathbf{\Gamma}\mathbf{\Gamma}^* - \tilde{\mathbf{\Gamma}}\tilde{\mathbf{\Gamma}}^* \quad (6.33)$$

où $\mathbf{\Gamma}$ est la matrice $M^2 \times L$ matrix dont les colonnes $\mathbf{\Gamma}_{i,j}$, $1 \leq i \leq j \leq M$ sont

$$\mathbf{\Gamma}_{i,i} = \mathbf{e}_i \otimes \mathbf{e}_i$$

et, pour $i < j$,

$$\mathbf{\Gamma}_{i,j} = \frac{1}{\sqrt{2}} (\mathbf{e}_i \otimes \mathbf{e}_j + \mathbf{e}_j \otimes \mathbf{e}_i)$$

alors que $\tilde{\mathbf{\Gamma}}$ est la matrice $M^2 \times M(M-1)/2$ dont les colonnes $\tilde{\mathbf{\Gamma}}_{i,j}$, $1 \leq i < j \leq M$ sont

$$\tilde{\mathbf{\Gamma}}_{i,j} = \frac{1}{\sqrt{2}} (\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i)$$

Les vecteurs $(\mathbf{e}_i)_{i=1,\dots,M}$ représentent la base canonique de \mathbb{C}^M . Il est facile de vérifier que $\mathbf{\Gamma}$ vérifie $\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathbf{I}_L$ et que $\mathbf{\Gamma}\mathbf{\Gamma}^* = \frac{1}{2}(\mathbf{I} + \mathbf{\Pi})$. Par conséquent, pour tous vecteurs de dimension M ($\mathbf{x}_1, \mathbf{x}_2$), on a l'identité suivante:

$$\mathbf{\Gamma}\mathbf{\Gamma}^*(\mathbf{x}_1 \otimes \mathbf{x}_2) = \frac{1}{2} (\mathbf{x}_1 \otimes \mathbf{x}_2 + \mathbf{x}_2 \otimes \mathbf{x}_1) \quad (6.34)$$

Ecrivons $\hat{\mathbf{C}}$ sous la forme

$$\hat{\mathbf{C}} = \hat{\mathbf{E}} - \hat{\mathbf{D}} \quad (6.35)$$

où $\hat{\mathbf{E}}$ et $\hat{\mathbf{D}}$ représentent les premier et deuxième termes du second membre de (6.30). Il est clair que $\mathbf{\Gamma}\mathbf{\Gamma}^*\hat{\mathbf{E}} = \hat{\mathbf{E}}$ et que $\hat{\mathbf{D}}$ peut se mettre sous la forme

$$\hat{\mathbf{D}} = 2 \frac{1}{N^2} \sum_{m,n} (\mathbf{y}_m \otimes \mathbf{y}_n) (\mathbf{y}_m \otimes \mathbf{y}_n)^* \mathbf{\Gamma}\mathbf{\Gamma}^* \quad (6.36)$$

Si l'on désigne par $\hat{\mathbf{R}}$ matrice de covariance empirique définie par

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^* \quad (6.37)$$

alors, on vérifie que $\hat{\mathbf{D}}$ est donné par

$$\hat{\mathbf{D}} = 2 (\hat{\mathbf{R}} \otimes \hat{\mathbf{R}}) \mathbf{\Gamma}\mathbf{\Gamma}^* \quad (6.38)$$

Par conséquent, la matrice $\hat{\mathbf{D}}$ est de rang $L = M(M+1)/2$. On peut également montrer facilement que si $(\hat{\mu}_i)_{i=1,\dots,\min(M,N)}$ sont les valeurs propres non nulles de $\hat{\mathbf{R}}$, alors les valeurs propres non nulles de $\hat{\mathbf{D}}$ sont les $(2\hat{\mu}_i\hat{\mu}_j)_{i \leq j}$.

6.3 Le cas où M et N tendent vers l'infini au même rythme.

Dans le premier rapport bibliographique, nous nous sommes intéressés au cas où M et N tendent vers l'infini au même rythme, i.e. que $\frac{M}{N} \rightarrow c$. Dans ce contexte, le rang de la matrice $\hat{\mathbf{E}}$, c'est-à-dire N , est beaucoup plus faible que sa dimension M^2 . En utilisant les résultats de [4], nous avons établi que les N des valeurs propres non nulles de $\hat{\mathbf{E}}$ se mettent sous la forme $2\sigma^4 N + o(N)$. De plus, si l'on suppose que $c < 1$ pour simplifier la présentation des résultats, la matrice $\hat{\mathbf{D}}$ est de rang $M(M+1)/2$, et ses valeurs propres non nulles sont presque sûrement localisées dans un voisinage de $[-2\sigma^4(1+\sqrt{c})^4, -2\sigma^4(1-\sqrt{c})^4]$. Les deux termes de (6.30) se comportent donc de façon bien différente puisque toutes les valeurs propres non nulles de $\hat{\mathbf{E}}$ tendent vers ∞ tandis que $\hat{\mathbf{D}}$ est bornée. Les inégalités de Weyl permettent de déduire de tout cela que les N plus grandes valeurs propres de $\hat{\mathbf{C}}$ sont toujours de la forme $2\sigma^4 N + o(N)$ et que les $M(M+1)/2 - N$ valeurs propres non nulles restantes se trouvent au voisinage de $[-2\sigma^4(1+\sqrt{c})^4, -2\sigma^4(1-\sqrt{c})^4]$. Par conséquent, le comportement de $\hat{\mathbf{C}}$ est bien différent de celui de \mathbf{C} , qui, en l'absence de bruit, vaut 0. En présence de

signal, il est très vraisemblable que le comportement de $\hat{\mathbf{E}}$ soit similaire, de sorte que les termes dus aux signaux utiles se trouvent complètement cachés par les contributions de \mathbf{v}_n . La matrice $\hat{\mathbf{D}}$ reste elle toujours bornée, avec des valeurs propres qui peuvent s'échapper de $[-2\sigma^4(1 + \sqrt{c})^4, -2\sigma^4(1 - \sqrt{c})^4]$. Toute l'information sur les signaux utiles se trouve donc dans $\hat{\mathbf{D}}$, qui porte essentiellement l'information contenue dans les statistiques du second ordre, insuffisante pour réaliser la séparation des sources. Cette discussion permet de penser que le régime dans lequel $\frac{M}{N} \rightarrow c$ ne permet pas aux approches basées sur $\hat{\mathbf{C}}$ de fonctionner. Il convient donc d'étudier des situations dans lesquelles le nombre d'observations est d'un ordre plus grand que le nombre de capteurs.

6.4 Le cas où N est de l'ordre de M^2 .

Nous étudions donc le cas dans lequel N est de l'ordre de M^2 , ou de façon équivalente de $L = M(M + 1)/2$. Nous désignons par c_N le rapport $c_N = \frac{L}{N}$, et supposons que

M et N tendent vers l'infini de telle façon que $c_N \rightarrow c$

Dans ce régime, la matrice de covariance empirique $\hat{\mathbf{R}}$ converge au sens de la norme spectrale vers $\sigma^2 \mathbf{I}$ car le rapport $\frac{M}{N} \rightarrow 0$. Dans ces conditions, la matrice $\hat{\mathbf{D}}$ tend elle même en norme vers $2\sigma^4 \mathbf{\Gamma} \mathbf{\Gamma}^*$. Nous posons dans la suite

$$\boldsymbol{\xi}_n = \frac{1}{N^{1/2}} \mathbf{\Gamma}^* (\mathbf{v}_n \otimes \mathbf{v}_n) \quad (6.39)$$

et $\boldsymbol{\Sigma} = (\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N)$. Il est clair que $\mathbf{\Gamma} \boldsymbol{\xi}_n = \frac{1}{N^{1/2}} (\mathbf{v}_n \otimes \mathbf{v}_n)$, et que par conséquent,

$$\hat{\mathbf{E}} = \mathbf{\Gamma} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^* \mathbf{\Gamma}^*$$

Cela implique que la matrice $\hat{\mathbf{C}}$ a le même comportement en norme que $\mathbf{\Gamma} (\boldsymbol{\Sigma} \boldsymbol{\Sigma}^* - 2\sigma^4 \mathbf{I}_L) \mathbf{\Gamma}^*$. Afin d'étudier les valeurs propres de $\hat{\mathbf{C}}$, il suffit donc de s'intéresser au comportement de la matrice $\boldsymbol{\Sigma} \boldsymbol{\Sigma}^*$. Les vecteurs colonnes de $\boldsymbol{\Sigma}$ sont indépendants identiquement distribués, et il est facile de vérifier que $\mathbb{E}(\boldsymbol{\xi}_n \boldsymbol{\xi}_n^*) = \frac{2\sigma^4}{N} \mathbf{I}_L$. Par ailleurs, on peut montrer le résultat suivant.

Proposition 1 *Soit \mathbf{B} une matrice déterministe $L \times L$. Alors, pour entier p , on a l'inégalité suivante*

$$\mathbb{E} \left| \boldsymbol{\xi}_j^* \mathbf{B} \boldsymbol{\xi}_j - 2\sigma^4 c_N \frac{1}{L} \text{Tr}(\mathbf{B}) \right|^p \leq \frac{C}{N^{p/4}} \|\mathbf{B}\|^{p/4} \quad (6.40)$$

où C est une constante.

La Proposition 1 implique donc que les formes quadratiques $\boldsymbol{\xi}_j^* \mathbf{B} \boldsymbol{\xi}_j$ se concentrent autour de $2\sigma^4 c_N \frac{1}{L} \text{Tr}(\mathbf{B})$. Les résultats de [9] permettent d'en déduire que la distribution empirique des valeurs propres de la matrice $\boldsymbol{\Sigma} \boldsymbol{\Sigma}^*$ tend vers la distribution de Marcenko-Pastur de paramètres $2\sigma^4$ et c . Afin de pouvoir par la suite traiter le cas où des sources sont présentes, il convient également de s'intéresser à la localisation presque sûre des valeurs propres de $\boldsymbol{\Sigma} \boldsymbol{\Sigma}^*$, et de vérifier si presque sûrement, à partir d'un certain rang, les $\min(L, N)$ valeurs propres non nulles de $\boldsymbol{\Sigma} \boldsymbol{\Sigma}^*$ se trouvent localisées au voisinage de l'intervalle $[2\sigma^4(1 - \sqrt{c})^2, 2\sigma^4(1 + \sqrt{c})^2]$. Plus formellement, nous avons été en mesure d'établir le résultat suivant.

Théorème 8 *Pour tout $\epsilon > 0$, presque sûrement, il existe un entier $N_0 \in \mathbb{N}$ tel que pour tout $N > N_0$, les $\min(L, N)$ valeurs propres non nulles de $\boldsymbol{\Sigma} \boldsymbol{\Sigma}^*$ appartiennent à $[2\sigma^4(1 - \sqrt{c})^2 - \epsilon, 2\sigma^4(1 + \sqrt{c})^2 + \epsilon]$.*

Pour établir le Théorème, nous avons utilisé l'approche de [5] et [10]. Nous préférons ne pas développer plus ce point technique qui nécessite d'effectuer des calculs très laborieux.

Le théorème 8 ainsi que la convergence en norme de $\hat{\mathbf{D}}$ impliquent que les valeurs propres de $\hat{\mathbf{C}}$ se situent au voisinage de l'intervalle $[2\sigma^4(c - 2\sqrt{c}), 2\sigma^4(c + 2\sqrt{c})]$. Si $c \rightarrow 0$, la largeur de l'intervalle tend vers 0, et on retrouve la situation traditionnelle dans laquelle $\hat{\mathbf{C}}$ tend vers 0. La présente analyse a donc le mérite de montrer que pour que $\hat{\mathbf{C}}$ tende vers

0, il faut que $\frac{M^2}{N} \rightarrow 0$. Par ailleurs, si cette condition n'est pas respectée, nous avons également pu quantifier de façon précise l'étalement des valeurs propres.

En termes de perspective, outre la prise en compte de bruits plus généraux, nous allons considérer le cas où un nombre fini de sources bande étroite sont également présentes. Afin d'expliciter le type de résultat que nous souhaitons établir, considérons le cas où $K = 1$. Dans ce contexte, la contribution de la source à l'observation \mathbf{y}_n se met sous la forme $\mathbf{a}_N s_n$ où $(s_n)_{n \in \mathbb{Z}}$ est un signal déterministe et où le vecteur de dimension M \mathbf{a}_N est déterministe. Notons \bar{s}_2 et \bar{s}_4 les termes définis par

$$\bar{s}_i = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=1}^N |s_n|^i$$

pour $i = 2, 4$, et appelons ρ la limite quand N tend vers l'infini de $\|\mathbf{a}_N\|^2$. On désigne par $\psi(\lambda)$ la fonction définie par

$$\psi(\lambda) = \frac{(\lambda + 2\sigma^4 c)(\lambda + 2\sigma^4)}{\lambda}$$

Alors, nous avons un bon nombre de raisons de penser que les valeurs propres de $\Sigma \Sigma^*$ se comportent comme celles d'un modèle information plus bruit défini par

$$(\mathbf{H} + \mathbf{W})(\mathbf{H} + \mathbf{W})^*$$

où \mathbf{H} est la matrice définie par

$$\mathbf{H} = \mathbf{\Gamma}^* [(4\sigma^2 \bar{s}_2 \mathbf{I}_M + \bar{s}_4 \mathbf{a} \mathbf{a}^*) \otimes \mathbf{a} \mathbf{a}^*] \mathbf{\Gamma}$$

et où les éléments de \mathbf{W} sont des variables aléatoires gaussiennes complexes indépendantes et identiquement distribuées de variance $\frac{2\sigma^4}{N}$. Les valeurs propres non nulles de \mathbf{H} convergent vers $\lambda_1 = \rho^2 \bar{s}_4 + 4\sigma^2 \rho \bar{s}_2$ avec multiplicité 1 et $\lambda_2 = 2\sigma^2 \rho \bar{s}_2$ avec multiplicité $M - 1$, nous conjecturons que les valeurs propres de $\Sigma \Sigma^*$, donc celles de $\hat{\mathbf{E}}$, ont le comportement suivant.

- Si $\lambda_1 \leq 2\sigma^4 \sqrt{c}$, la plus grande valeur propre de $\hat{\mathbf{E}}$ converge vers $2\sigma^4(1 + \sqrt{c})^2$ comme dans le cas où il n'y a pas de source.
- Si $\lambda_1 > 2\sigma^4 \sqrt{c}$ et si $\lambda_2 \leq 2\sigma^4 \sqrt{c}$, la plus grande valeur propre de $\hat{\mathbf{E}}$ converge vers $\psi(\lambda_1)$ et les autres valeurs propres sont situées au voisinage de $[2\sigma^4(1 - \sqrt{c})^2, 2\sigma^4(1 + \sqrt{c})^2]$
- Si $\lambda_2 > 2\sigma^4 \sqrt{c}$, la plus grande valeur propre de $\hat{\mathbf{E}}$ converge vers $\psi(\lambda_1)$, les $M - 1$ suivantes convergent toutes vers la valeur $\psi(\lambda_2)$, et les autres sont au voisinage de l'intervalle $[2\sigma^4(1 - \sqrt{c})^2, 2\sigma^4(1 + \sqrt{c})^2]$.

De surcroît, les vecteurs propres correspondant aux valeurs propres dépassant le seuil $2\sigma^4(1 + \sqrt{c})^2$ devraient pouvoir également être quantifiés, et rapprochés des vecteurs propres de la matrice \mathbf{H} . On peut ainsi espérer avoir une idée assez précise des valeurs propres de $\hat{\mathbf{C}}$ s'échappant de l'intervalle $[2\sigma^4(c - 2\sqrt{c}), 2\sigma^4(c + 2\sqrt{c})]$ et des vecteurs propres associés. Ces informations devraient permettre d'adapter les algorithmes de séparation de sources basés sur $\hat{\mathbf{C}}$.

Références

- [1] L. Albéra, A. Ferréol, P. Chevalier, P. Comon, "ICAR, a tool for blind source separation using fourth order statistics only", *IEEE Transactions On Signal Processing*, vol. 53, n°10, part 1, pp. 3633-3643, October 2005.
- [2] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.*, 26(1):316–345, 1998.
- [3] 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.

- [4] N. El-Karoui, "The spectrum of kernel random matrices", *Annals of Statistics*, vol. 38, no. 1, pp. 1-50, January 2010.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] X. Mestre, M.A Lagunas, "Modified subspace algorithms for DoA estimation with large arrays", *IEEE Transactions on Signal Processing*, vol. 56, no.2, pp. 598-614, 2008.
- [9] A. Pajor, L.A. Pastur, "On the Limiting Empirical Measure of the sum of rank one matrices with log-concave distribution, *Studia Math.* 195 (2009), 11-29.
- [10] H. Schultz, "Non commutative polynomials of independent Gaussian random matrices", *Probab. Theory Relat. Fields* 131, 261-309 (2005)
- [11] J. W. Silverstein and Sang-Il Choi. Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):295–309, 1995.
- [12] Pascal Vallet, Philippe Loubaton, Xavier Mestre, "Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case", *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1043-1068, February 2012.
- [13] P. Vallet, P. Loubaton, "Toeplitz rectification and DoA estimation with MUSIC, in Proceedings ICASSP 2014, pp. 2237-2241, Florence, May 2014.
- [14] P. Vallet, P. Loubaton, X. Mestre "A statistical comparison between MUSIC and G-MUSIC", soumis à ICASSP 2015.
- [15] L. Zhang. *Spectral Analysis of Large Random Matrices*. PhD thesis, National University of Singapore, 2006.
- [16] J. Vinogradova, R. Couillet and W. Hachem, "Statistical Inference in Large Antenna Arrays under Unknown Noise Pattern", *IEEE Transactions on Signal Processing*, Vol. 61, no. 22, pp. 5633-5645, 2013.
- [17] J. Vinogradova, R. Couillet, and W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection", submitted to *IEEE Transactions on Signal Processing* in March 2014. Revised in July 2014.
- [18] J. Baik, G. Ben Arous and S. Péché, "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices", *Ann. Probab.*, 33(5):1643–1697, 2005.
- [19] N. El Karoui, "Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices", *Ann. Probab.*, 35(2):663–714, 2007.
- [20] W. Hachem, A. Hardy and J. Najim, "Large Complex Correlated Wishart Matrices: Fluctuations and Asymptotic Independence at the Edges". Submitted, September 2014.
- [21] F. Bornemann "Asymptotic independence of the extreme eigenvalues of Gaussian unitary ensemble", *J. Math. Phys.*, 51(2):023514, 8, 2010.

- [22] W. Hachem and R. Couillet, "Analysis of the limiting spectral measure of large random matrices of the separable covariance type", accepted for publication in *Random Matrices: Theory and Applications*, in October 2014.
- [23] A. Boutet de Monvel, A. Khorunzhy, and V. Vasilchuk, "Limiting eigenvalue distribution of random matrices with correlated entries", *Markov Process. Related Fields*, 2(4):607–636, 1996.
- [24] C. Butucea, R. Zgheib, "Sharp minimax tests for large covariance matrices", arxiv:1409.1429
- [25] P. Bianchi, M. Debbah, M. Maïda and J. Najim, "Performance of Statistical Tests for Single Source Detection using Random Matrix Theory", *IEEE Inf. Theory*, Vol. 57 (4), april 2011, 2400–2419.