

Projet ANR SESAME
Rapport sur le sous-projet 1 (WP1) :
Revue de la littérature

Janvier 2009

Ce rapport passe en revue certains résultats existants sur lesquels s'appuieront les travaux du projet SESAME. Deux classes de problèmes sont à l'étude dans SESAME :

1. L'estimation consistante de certains paramètres d'un modèle de signaux à partir de la matrice de covariance empirique. Ce problème est décliné en deux sous-problèmes : l'estimation consistante dans le cadre des communications CDMA (« Code Division Multiple Access ») qui fait l'objet du WP2, et l'estimation consistante pour les modèles dits « Information plus Bruit » (WP3).
2. L'amélioration de la matrice de covariance empirique par la prise en compte d'une structure inhérente à la « vraie » matrice de covariance (WP4).

Ces deux classes de problèmes seront traitées dans ce document dans les chapitres 1 et 2 respectivement.

Chapitre 1

Matrices de covariance empiriques et estimation consistante

1.1 Introduction

Le comportement asymptotique des mesures spectrales de deux modèles de matrices de Gram $\Sigma\Sigma^*$ présents dans les sous-projets WP2 et WP3 [23] seront traités dans ce chapitre. Ces deux modèles sont les suivants :

- Le modèle que nous qualifions de « modèle centré »

$$\Sigma = \frac{1}{\sqrt{K}}\mathbf{H}\mathbf{X}\mathbf{P}^{1/2}$$

où Σ est de taille $M \times K$, les entiers M et K tendent vers l'infini à la même vitesse,

\mathbf{H} est une matrice déterministe $M \times M$,

\mathbf{X} est une matrice aléatoire $M \times K$ dont les éléments X_{mk} sont indépendants et identiquement distribués (iid), $\mathbb{E}X_{11} = 0$ et $\mathbb{E}|X_{11}|^2 = 1$,

\mathbf{P} est une matrice déterministe diagonale positive $K \times K$.

- Le modèle dit « Information plus bruit » où Σ (toujours de dimensions $M \times K$) s'écrit

$$\Sigma = \mathbf{A} + \frac{1}{\sqrt{K}}\mathbf{X}$$

où

\mathbf{A} est une matrice déterministe $M \times K$,

\mathbf{X} est toujours une matrice aléatoire $M \times K$ dont les éléments X_{mk} sont iid, $\mathbb{E}X_{11} = 0$ et $\mathbb{E}|X_{11}|^2 = 1$.

Nous commencerons par rappeler la définition et les propriétés importantes de la transformée de Stieltjes, transformée qui joue un rôle fondamentale en théorie des matrices aléatoires. Ensuite, pour le modèle centré, nous passerons en revue les résultats relatifs à l'approximation déterministe de la mesure spectrale de $\Sigma\Sigma^*$ par

le biais de la transformée de Stieltjes, les propriétés de la mesure limite, le comportement des valeurs propres extrêmes et la convergence par l'intermédiaire d'un Théorème de la Limite Centrale (TLC). Toujours dans le cas du modèle centré, nous exposerons ensuite la technique d'estimation consistante développée par X. Mestre dans [21].

Nous terminons le chapitre par l'exposé des résultats connus qui concernent la mesure spectrale de $\Sigma\Sigma^*$ pour le modèle « Information plus bruit ». Les techniques d'estimation consistante pour ce modèle (analogues aux travaux de Mestre pour le modèle centré) sont à l'étude.

1.2 La transformée de Stieltjes

Comme annoncé plus haut, l'essentiel de ce document concerne l'étude du comportement asymptotique de la mesure spectrale d'une matrice aléatoire quand les dimensions de cette matrice convergent vers l'infini à la même vitesse. Rappelons que la mesure spectrale d'une matrice hermitienne \mathbf{A} de dimensions $N \times N$ est la mesure de probabilité sur \mathbb{R}

$$\mu(dx) = \frac{1}{N} \sum_{n=1}^N \delta_{\lambda_n}(dx)$$

où $\{\lambda_1, \dots, \lambda_N\}$ est l'ensemble des valeurs propres de \mathbf{A} . Afin d'étudier le comportement d'une mesure spectrale, un rôle fondamental sera joué par la transformée de Stieltjes dont l'utilisation en théorie des matrices aléatoires remonte à l'article fondateur de Marchenko et de Pastur [20]. Soit μ une mesure de probabilité sur \mathbb{R} . La transformée de Stieltjes $m_\mu(z)$ de cette mesure est la fonction complexe

$$m_\mu(z) = \int \frac{1}{\lambda - z} \mu(d\lambda) . \quad (1.1)$$

Cette fonction est analytique sur $\mathbb{C}_+ = \{z \in \mathbb{C}, \Im(z) > 0\}$. La formule d'inversion

$$\mu([a, b]) = \lim_{v \downarrow 0} \frac{1}{\pi} \int_a^b \Im(m_\mu(u + iv)) du \quad (1.2)$$

où a et b sont des points de continuité de μ nous montre que μ est associée à sa transformée de Stieltjes d'une manière biunivoque.

Par ailleurs, la convergence faible $\mu_N \implies \mu$ d'une suite de mesures de probabilité μ_N vers μ équivaut à la convergence ponctuelle de leurs transformées de Stieltjes :

$$(\mu_N \implies \mu) \Leftrightarrow (m_{\mu_N}(z) \rightarrow m_\mu(z), z \in \mathbb{C}_+) . \quad (1.3)$$

Le sens \implies de cette équivalence se déduit de la définition (1.1) de la transformée de Stieltjes et le sens \Leftarrow de (1.2).

Quand la mesure de probabilité en question est la mesure spectrale d'une matrice hermitienne \mathbf{A} , sa transformée de Stieltjes est intimement associée à la résolvante de cette matrice, qui est la fonction matricielle $\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I})^{-1}$. En effet, soit

$\mathbf{A} = \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_N) \mathbf{U}^*$ une décomposition spectrale de \mathbf{A} et soit $\mu = \frac{1}{N} \sum \delta_{\lambda_i}$ sa mesure spectrale. Nous avons

$$\begin{aligned} \frac{1}{N} \text{tr}(\mathbf{Q}(z)) &= \frac{1}{N} \text{tr}(\mathbf{A} - z\mathbf{I}_N)^{-1} \\ &= \frac{1}{N} \text{tr} \left(\mathbf{U} \begin{bmatrix} \frac{1}{\lambda_1 - z} & & \\ & \ddots & \\ & & \frac{1}{\lambda_N - z} \end{bmatrix} \mathbf{U}^* \right) \\ &= \frac{1}{N} \sum_{n=1}^N \frac{1}{\lambda_n - z} \\ &= \int \frac{1}{t - z} d\mu(t) \\ &= m_\mu(z) . \end{aligned}$$

Ainsi, la trace normalisée de la résolvante de \mathbf{A} est liée d'une manière biunivoque à la mesure spectrale de cette matrice.

Afin d'étudier le comportement asymptotique de la mesure spectrale μ_N d'une matrice aléatoire, l'idée est donc d'étudier le comportement asymptotique de la trace normalisée de la résolvante. Dans certains cas, la mesure aléatoire μ_N converge faiblement (disons, au sens presque sûr) vers une mesure de probabilité déterministe μ . Afin d'établir cette convergence, grâce à (1.3) il suffit de prouver que la trace normalisée de la résolvante converge vers une fonction déterministe qui appartienne à la classe des transformées de Stieltjes de mesures de probabilité. Plus généralement, quand la convergence faible presque sûre $\mu_N \implies \mu$ n'est pas garantie, il est souvent possible de mettre en évidence l'existence d'une suite de mesures déterministes π_N qui constitue une approximation μ_N dans le sens où $m_{\mu_N}(z) - m_{\pi_N}(z) \rightarrow 0$ presque sûrement dans le régime des grandes dimensions. Par le biais de $m_{\pi_N}(z)$ on obtient des informations sur π_N (densité, support, ...). Une étude « fine » de la variable aléatoire $m_{\mu_N}(z) - m_{\pi_N}(z)$ permet d'étudier la convergence, le comportement des valeurs propres extrêmes, un TLC, etc.

1.3 Mesure spectrale du modèle centré

Le but de ce paragraphe est d'étudier le comportement de la mesure spectrale de la matrice de Gram $\Sigma \Sigma^*$ pour le modèle centré quand M et K tendent vers l'infini à la même vitesse.

1.3.1 Convergence de la mesure spectrale

Ce problème a été étudié dans de nombreux articles, souvent dans le cas où $\mathbf{P} = \mathbf{I}_K$ ou dans le cas où les éléments de \mathbf{X} sont gaussiens (citons sans être exhaustifs [12, 14, 18, 20, 24, 25, 27, 30]). Nous avons le théorème suivant :

Théorème 1. *Pour $K = 1, 2, \dots$,*

- Soit $M = M(K)$ une suite d'entiers telle que

$$0 < \liminf_K \frac{M(K)}{K} \leq \limsup_K \frac{M(K)}{K} < \infty .$$

- Soit $\mathbf{X}_K = [X_{mk}^{(K)}]$ une matrice aléatoire $M \times K$ dont les éléments $X_{mk}^{(K)} \in \mathbb{C}$ constituent un tableau $M \times K$ d'éléments iid, $\mathbb{E}X_{11}^{(K)} = 0$ et $\mathbb{E}|X_{11}^{(K)}|^2 = 1$.
- Soit \mathbf{H}_K une suite de matrices déterministes $M \times M$ dont la norme spectrale satisfait $\sup_K \|\mathbf{H}_K\| < \infty$. On désigne par ν_K la mesure spectrale de $\mathbf{H}_K \mathbf{H}_K^*$.
- Soit \mathbf{P}_K une suite de matrices déterministes diagonales positives $K \times K$ telles que $\sup_K \|\mathbf{P}_K\| < \infty$. On note $\tilde{\nu}_K$ la mesure spectrale de \mathbf{P}_K .

Alors,

- Pour tout K , le système d'équations

$$\delta_K(z) = \frac{M}{K} \int \frac{t}{-z(1 + \tilde{\delta}_K(z)t)} \nu_K(dt) \quad (1.4)$$

$$\tilde{\delta}_K(z) = \int \frac{t}{-z(1 + \delta_K(z)t)} \tilde{\nu}_K(dt) \quad (1.5)$$

admet une solution unique $(\delta_K, \tilde{\delta}_K) \in \mathbb{C}_+^2$ pour $z \in \mathbb{C}_+$. Par ailleurs, la fonction complexe

$$p_K(z) = \int \frac{1}{-z(1 + \tilde{\delta}_K(z)t)} \nu_K(dt) \quad (1.6)$$

est la transformée de Stieltjes d'une mesure de probabilité π_K portée par $\mathbb{R}_+ = [0, \infty)$.

- Soit Σ_K la matrice aléatoire $M \times K$

$$\Sigma_K = \frac{1}{\sqrt{K}} \mathbf{H}_K \mathbf{X}_K \mathbf{P}_K^{1/2} .$$

Pour tout $z \in \mathbb{C}_+$, la transformée de Stieltjes $m_K(z) = \frac{1}{M} \text{tr}(\Sigma_K \Sigma_K^* - z\mathbf{I})^{-1}$ associée à la mesure spectrale μ_K de $\Sigma_K \Sigma_K^*$ satisfait

$$m_K(z) - p_K(z) \xrightarrow{K \rightarrow \infty} 0 \quad \text{presque sûrement.}$$

- Pour toute fonction réelle φ continue et bornée, nous avons

$$\int \varphi(\lambda) \mu_K(d\lambda) - \int \varphi(\lambda) \pi_K(d\lambda) \xrightarrow{K \rightarrow \infty} 0 \quad \text{presque sûrement.}$$

Il est utile de considérer également la mesure spectrale $\tilde{\mu}_K$ de $\Sigma^* \Sigma$ dont la transformée de Stieltjes sera notée $\tilde{m}_K(z)$. L'homologue $\tilde{\pi}_K$ de π_K a comme transformée de Stieltjes la fonction $\tilde{p}_K(z)$ donnée par

$$\tilde{p}_K(z) = \int \frac{1}{-z(1 + \delta_K(z)t)} \tilde{\nu}_K(dt) . \quad (1.7)$$

En considérant les décompositions spectrales de $\Sigma\Sigma^*$ et de $\Sigma^*\Sigma$, on vérifie facilement que $\tilde{m}_K(z) = c_K m_K(z) + (c_K - 1) \frac{1}{z}$ où on pose $c_K = M/K$. On a aussi

$$\tilde{p}_K(z) = c_K p_K(z) + (c_K - 1) \frac{1}{z}. \quad (1.8)$$

Dans le cas où c_K converge et où ν_K et $\tilde{\nu}_K$ convergent faiblement, la mesure spectrale ν_K converge vers une mesure limite :

Corollaire 1. *Supposons que $M/K \rightarrow c > 0$, que $\nu_K \Rightarrow \mu$ et que $\tilde{\nu}_K \Rightarrow \tilde{\mu}$ quand $K \rightarrow \infty$. Alors, avec probabilité un, la mesure spectrale μ_K converge faiblement vers une mesure de probabilité μ portée par \mathbb{R}_+ . La transformée de Stieltjes $m(z)$ de μ est donnée par*

$$m(z) = \int \frac{1}{-z(1 + \tilde{\delta}(z)t)} \mu(dt)$$

où pour tout $z \in \mathbb{C}_+$ le couple $(\delta(z), \tilde{\delta}(z))$ est l'unique solution dans \mathbb{C}_+^2 du système d'équations

$$\delta(z) = c \int \frac{t}{-z(1 + \tilde{\delta}(z)t)} \nu(dt), \quad \tilde{\delta}(z) = \int \frac{t}{-z(1 + \delta(z)t)} \tilde{\nu}(dt).$$

Dans ces conditions, nous avons bien sûr $\pi_K \Rightarrow \mu$ et $p_K(z) \rightarrow m(z)$.

Dans la suite, nous nous limiterons au cas où $\mathbf{P} = \mathbf{I}$. Notons que le cas où $\mathbf{H}\mathbf{H}^* = \mathbf{I}$ se traite d'une manière similaire car les spectres des matrices $\Sigma\Sigma^*$ et $\Sigma^*\Sigma$ ne diffèrent que par $|M - K|$ valeurs propres nulles.

Dans le cas où $\mathbf{P} = \mathbf{I}$, la mesure $\tilde{\nu}_K$ coïncide avec la mesure de Dirac en 1. En revenant aux équations (1.5), (1.7) et (1.8), la transformée de Stieltjes $\tilde{p}_K(z)$ de $\tilde{\pi}_K$ satisfait

$$\tilde{p}_K(z) = \tilde{\delta}_K(z) = \frac{-1}{z(1 + \delta_K(z))} = c_K p + (c_K - 1) \frac{1}{z}$$

et la transformée de Stieltjes $p_K(z)$ s'écrit (cf. (1.6))

$$p_K(z) = \int \frac{1}{-z - tz\tilde{p}_K(z)} \nu_K(dt) = \int \frac{1}{-z + t(1 - c_K - c_K z p_K(z))} \nu_K(dt). \quad (1.9)$$

Théorème 2. *Supposons que les hypothèses du théorème 1 soient satisfaites, et que par ailleurs, $\mathbf{P} = \mathbf{I}$. Pour tout $z \in \mathbb{C}_+$, $p = p_K(z)$ est la solution unique de l'équation (1.9) dans l'ensemble $\{p \in \mathbb{C} : -(1 - c_K)/z + c_K p \in \mathbb{C}_+\}$. Pour tout $z \in \mathbb{C}_+$, la transformée de Stieltjes $m_K(z) = \frac{1}{M} \text{tr}(\Sigma_K \Sigma_K^* - z\mathbf{I})^{-1}$ associée à la mesure spectrale μ_K de $\Sigma_K \Sigma_K^*$ satisfait*

$$m_K(z) - p_K(z) \xrightarrow{K \rightarrow \infty} 0 \quad \text{presque sûrement.}$$

Il sera également utile de caractériser l'approximation déterministe de la mesure spectrale de $\Sigma\Sigma^*$ par l'intermédiaire de \tilde{p}_K . Cette fonction est en effet la solution unique dans \mathbb{C}_+ de l'équation

$$\tilde{p} = \frac{1}{-z + c_K \int \frac{t}{1 + \tilde{p}t} \nu_K(dt)} \quad (1.10)$$

pour $z \in \mathbb{C}_+$. Il s'ensuit que \tilde{p} possède en tant que fonction de z une réciproque sur \mathbb{C}_+ donnée par

$$z(\tilde{p}) = -\frac{1}{\tilde{p}} + c_K \int \frac{t}{1 + \tilde{p}t} \nu_K(dt) \quad (1.11)$$

1.3.2 Propriétés des approximations déterministes $\pi_K, \tilde{\pi}_K$

Elles ont été étudiées par Silverstein et Choi [26] dans le cas où $\mathbf{P} = \mathbf{I}$. Rappelons qu'une mesure de probabilité se déduit de sa transformée de Stieltjes par la formule d'inversion (1.2). En étudiant le comportement au voisinage de l'axe des réels de la solution de (1.10), les auteurs de [26] ont établi le théorème suivant :

Theorème 3. *Pour tout $x \in \mathbb{R}^* = \mathbb{R} - \{0\}$, la limite $\lim_{z \in \mathbb{C}_+ \rightarrow x} \tilde{p}(z) \equiv \tilde{p}(x)$ existe. La fonction $\tilde{p}(x)$ est continue sur \mathbb{R}^* . La mesure de probabilité $\tilde{\pi}$ dont la transformée de Stieltjes est $\tilde{p}(z)$ possède une densité continue f sur \mathbb{R}^* donnée par $f(x) = \frac{1}{\pi} \Im \tilde{p}(x)$. Cette densité est une fonction analytique en tout $x \neq 0$ pour lequel $f(x) > 0$. Pour ces valeurs de x , $\pi f(x)$ est la partie imaginaire de l'unique $\tilde{p} \in \mathbb{C}_+$ qui satisfait*

$$x = -\frac{1}{\tilde{p}} + c_K \int \frac{t}{1 + \tilde{p}t} \nu_K(dt) . \quad (1.12)$$

L'article [26] fournit également un moyen de déterminer le support $\text{supp}(\tilde{\pi})$ de $\tilde{\pi}$ dans \mathbb{R}^* . En revenant à la définition (1.1) de la transformée de Stieltjes d'une mesure de probabilité μ , on constate que sur les intervalles ouverts de $\mathbb{R} - \text{supp}(\mu)$, la fonction $m_\mu(x)$ est définie, continue, réelle et croissante. Par conséquent sa réciproque existe sur ces intervalles et y est réelle et croissante. Les auteurs de [26] vont plus loin et prouvent qu'il est possible de *déterminer* $\text{supp}(\tilde{\pi})$ à partir des intervalles pour lesquels la réciproque donnée par (1.12) est croissante. D'une manière plus précise, le domaine de définition de $x(\tilde{p})$ est $\Delta = \{\tilde{p} \in \mathbb{R} : \tilde{p} \neq 0, -\tilde{p} \in \mathbb{R} - \text{supp}(\nu_K)\}$. Afin de déterminer $\text{supp}(\tilde{\pi})$, il suffit de tracer $x(\tilde{p})$ sur Δ , de repérer les plages où cette fonction est croissante et de supprimer tous les points en ordonnée correspondant à ces plages. Le support de $\tilde{\pi}$ (ou de π) sur \mathbb{R}^* est constitué des points restants sur l'axe des ordonnées. Cette même technique de détermination de ce support a été développée d'une manière indépendante dans [22]. Nous reviendrons plus loin sur ce sujet.

1.3.3 Valeurs propres extrêmes et théorème de séparation

Le comportement des valeurs propres dans l'intérieur de $\mathbb{R} - \cup_K \text{supp}(\pi_K)$ pour K grand a été caractérisé par Bai et Silverstein dans [2]. Toujours dans le cas où $\mathbf{P} = \mathbf{I}$ ces auteurs démontrent que pour tous K grands et avec probabilité un, $\Sigma \Sigma^*$ ne possède pas de valeur propre dans un intervalle fermé qui se trouve en dehors de l'union des supports des π_K [2] :

Theorème 4. *Supposons que*

- $M = M(K)$ satisfait $M(K)/K \rightarrow c > 0$.

- $\mathbf{X}_K = [X_{mk}^{(K)}]_{m,k=1}^{M,K}$ variables aléatoires complexes iid, $\mathbb{E}X_{11} = 0$, $\mathbb{E}|X_{11}|^2 = 1$ et $\mathbb{E}|X_{11}|^4 < \infty$.
- \mathbf{H}_K est une suite de matrices déterministes $M \times M$ telle que $\sup_K \|\mathbf{H}_K\| < \infty$ et la mesure spectrale ν_K de $\mathbf{H}_K \mathbf{H}_K^*$ converge faiblement vers une mesure de probabilité ν .

Pour $\Sigma_K = \frac{1}{\sqrt{K}} \mathbf{H}_K \mathbf{X}_K$, soit π_K l'approximant déterministe de la mesure spectrale de $\Sigma_K \Sigma_K^*$ déterminé par sa transformée de Stieltjes $p_K(z)$ donnée par (1.10). Soit μ la limite faible de la suite des mesures π_K . Soit $[a, b]$ où $a > 0$ un intervalle dont l'intersection avec $\text{supp}(\mu) \cup_K \text{supp}(\pi_K)$ pour tous K grands est vide. Soit \mathcal{A}_K l'événement « $\Sigma_K \Sigma_K^*$ possède au moins une valeur propre dans $[a, b]$ ». Alors

$$\mathbb{P} \left(\limsup_K \mathcal{A}_K \right) = 0$$

1.3.4 Théorème de la Limite Centrale

Dans les contextes des théorèmes 1 et 2, il est possible d'établir un TLC sur $\int \varphi(\lambda) (\mu_K(d\lambda) - \pi_K(d\lambda))$ pour certaines fonctions réelles φ . Ce problème a été abordé dans bon nombre d'articles, parmi lesquels nous citons [1, 4, 5, 15, 17]. En particulier, un TLC est établi dans [4] dans le cas où

1. Les X_{mk} sont iid réels, $\mathbb{E}X_{11} = 0$, $\mathbb{E}X_{11}^2 = 1$, $\mathbb{E}X_{11}^4 = 3$
ou
Les X_{mk} sont iid complexes, $\mathbb{E}X_{11} = \mathbb{E}X_{11}^2 = 0$, $\mathbb{E}|X_{11}|^2 = 1$, et $\mathbb{E}|X_{11}|^4 = 2$.
2. $\mathbf{P} = \mathbf{I}$.
3. La fonction réelle φ est analytique sur un intervalle ouvert suffisamment grand pour contenir

$$\left[\liminf_K \lambda_{\min}(\mathbf{H}\mathbf{H}^*) \lambda_{\min}^{\text{MP}}, \limsup_K \lambda_{\max}(\mathbf{H}\mathbf{H}^*) \lambda_{\max}^{\text{MP}} \right]$$

où $\lambda_{\min}(\mathbf{H}\mathbf{H}^*)$ (resp. $\lambda_{\max}(\mathbf{H}\mathbf{H}^*)$) est la valeur propre minimum (resp. maximum) de $\mathbf{H}\mathbf{H}^*$ et $\lambda_{\min}^{\text{MP}}$ (resp. $\lambda_{\max}^{\text{MP}}$) est la valeur propre minimum (resp. maximum) de la mesure spectrale limite de $\frac{1}{K} \mathbf{X}\mathbf{X}^*$, dite loi de Marchenko-Pastur.

Rappelons que $\mathbb{E}X_{11}^4 = 3$ (resp. $\mathbb{E}|X_{11}|^4 = 2$) si X_{11} est gaussienne réelle standard (resp. gaussienne complexe circulaire standard). Nous ramenons le lecteur à [4] pour l'énoncé complet du TLC.

1.4 Modèle centré et estimation consistante

1.4.1 Présentation du problème et contexte applicatif

Dans [21], X. Mestre considère également le cas où $\mathbf{P} = \mathbf{I}$ et pose

$$\mathbf{H}\mathbf{H}^* = \sum_{m=1}^{\bar{M}} \gamma_m \mathbf{\Pi}_m$$

où $\gamma_1 < \gamma_2 < \dots < \gamma_{\bar{M}}$ sont les valeurs propres distinctes de $\mathbf{H}\mathbf{H}^*$ au nombre (fixe) de \bar{M} et où les $\mathbf{\Pi}_m$ sont les projecteurs orthogonaux sur les sous-espaces propres de $\mathbf{H}\mathbf{H}^*$. Il considère le régime asymptotique $K \rightarrow \infty$, $M/K \rightarrow c > 0$ et suppose que la mesure spectrale ν_K de $\mathbf{H}\mathbf{H}^*$ converge vers une mesure de probabilité ν inconnue constituée d'une somme de \bar{M} mesures de Dirac en $\{\gamma_1, \dots, \gamma_{\bar{M}}\}$ pondérées par les poids $(I_1/M, \dots, I_{\bar{M}}/M)$ connus. Il se propose d'estimer d'une manière consistante les γ_m ainsi que des produits scalaires du type $\eta_m = \mathbf{a}^* \mathbf{\Pi}_m \mathbf{b}$ où \mathbf{a} et \mathbf{b} représentent des vecteurs déterministes connus.

Un grand nombre de problèmes d'estimation en traitement du signal et en théorie des communications entrent dans ce cadre. Considérons le modèle

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k, \quad k = 1, \dots, K$$

où la suite de vecteurs aléatoires $(\mathbf{y}_k)_{k=1}^K$ représente K échantillons d'un signal reçu multivarié de dimension M (par exemple, un signal radio reçu sur un réseau de M antennes) et $(\mathbf{x}_k)_{k=1}^K$ est une suite de vecteurs iid telle que \mathbf{x}_1 soit à éléments iid centrés de variance unité. La matrice de covariance du signal reçu $\mathbf{R} = \mathbb{E}\mathbf{y}_1\mathbf{y}_1^* = \mathbf{H}\mathbf{H}^*$ est en général inconnue, et est traditionnellement remplacée par la matrice de covariance empirique

$$\hat{\mathbf{R}}_K = \frac{1}{K} \sum_{k=1}^K \mathbf{y}_k \mathbf{y}_k^* = \sum_{i=1}^M \lambda_i \mathbf{v}_i \mathbf{v}_i^* \quad (1.13)$$

où $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_M$ sont les valeurs propres de $\hat{\mathbf{R}}_K$ et les \mathbf{v}_i sont les vecteurs propres qui leur sont associés. Si M est fixe, la loi des grands nombres nous dit que $\hat{\mathbf{R}}_K \xrightarrow[K \rightarrow \infty]{\text{p.s.}} \mathbf{R}$. Ainsi, l'estimée $\hat{\mathbf{R}}_K$ est une estimée consistante de \mathbf{R} pour $K \gg M$. En particulier,

$$\hat{\gamma}_m = \frac{1}{I_m} \sum_{i \in \mathcal{I}_m} \lambda_i \quad \text{et} \quad \hat{\mathbf{\Pi}}_m = \sum_{i \in \mathcal{I}_m} \mathbf{v}_i \mathbf{v}_i^* \quad (1.14)$$

où \mathcal{I}_m est l'ensemble des indices

$$\mathcal{I}_m = \left\{ \sum_{k=1}^{m-1} I_k + 1, \sum_{k=1}^{m-1} I_k + 2, \dots, \sum_{k=1}^m I_k \right\}$$

sont des estimées consistantes de γ_m et de $\mathbf{\Pi}_m$.

En revanche, dans le cas fréquent en pratique où K et M sont du même ordre de grandeur, l'estimée $\hat{\mathbf{R}}$ n'est plus consistante. L'idée est alors de recourir à la théorie des grandes matrices aléatoires et de considérer plutôt le régime où $K \rightarrow \infty$ et $M/K \rightarrow c > 0$. En posant $\mathbf{X}_K = [\mathbf{x}_1 \dots \mathbf{x}_K]$, nous avons $\hat{\mathbf{R}}_K = \mathbf{\Sigma}_K \mathbf{\Sigma}_K^*$ où $\mathbf{\Sigma}_K$ satisfait les hypothèses du théorème 2 ci-dessus. Dans ces conditions, il s'avère possible de construire des estimées consistantes d'objets qui dépendent de \mathbf{R} (valeurs propres ou produits scalaires du type $\eta_m = \mathbf{a}^* \mathbf{\Pi}_m \mathbf{b}$). De tels estimateurs sont proposés dans [21]. Ici, il convient de noter que certaines des idées développées dans [21] se retrouvent dans les travaux antérieurs de V. Girko [13].

L'idée générale en est la suivante : on peut vérifier sans difficulté que la valeur propre γ_m de $\mathbf{R} = \mathbf{H}\mathbf{H}^*$ s'écrit aussi

$$\gamma_m = \frac{1}{I_m} \frac{1}{2\pi} \oint_{\mathcal{C}_m^-} z \operatorname{tr}(\mathbf{R} - z\mathbf{I})^{-1} dz = \frac{M}{I_m} \frac{1}{2\pi} \oint_{\mathcal{C}_m^-} z m_{\nu_K}(z) dz \quad (1.15)$$

où \mathcal{C}_m^- est un contour orienté négativement à valeurs dans $\mathbb{C} - \{\gamma_1, \dots, \gamma_{\bar{M}}\}$ et qui entoure γ_m uniquement. Dans le régime asymptotique $K \rightarrow \infty$, $M/K \rightarrow c > 0$, on ne peut pas remplacer \mathbf{R} par $\widehat{\mathbf{R}}$ car $m_K(z)$ n'est pas une estimée consistante de $m_{\nu_K}(z)$. Par contre, il s'avère possible de réaliser un changement de variable qui permettra d'exprimer l'intégrale (1.15) en fonction de la transformée de Stieltjes $p_K(z)$ au lieu de l'exprimer en fonction de $m_{\nu_K}(z)$. Grâce au théorème 2, $m_K(z)$ constitue une estimée consistante de $p_K(z)$, chose qui permettra d'estimer γ_m d'une manière consistante. Le changement de variable visé requiert une connaissance précise du support de la mesure π_K de transformée de Stieltjes $p_K(z)$. Dans ce but, nous présenterons au paragraphe suivant des résultats qui concernent le support de π_K et qui viendront compléter ceux du paragraphe 1.3.2.

Avant d'aborder ce sujet, nous présentons les résultats analogues à (1.15) et au théorème 2 pour $\eta_m = \mathbf{a}^* \mathbf{\Pi}_m \mathbf{b}$. Cette quantité s'écrit

$$\eta_m = \frac{1}{2i\pi} \oint_{\mathcal{C}_m^-} \mathbf{a}^* (\mathbf{R} - z\mathbf{I})^{-1} \mathbf{b} dz \quad (1.16)$$

L'homologue de $m_K(z)$ pour η_m est la fonction aléatoire

$$q_K(z) = \mathbf{a}^* (\mathbf{\Sigma}_K \mathbf{\Sigma}_K^* - z\mathbf{I})^{-1} \mathbf{b}$$

qui dépend clairement des valeurs propres et des sous-espaces propres de $\widehat{\mathbf{R}}_K$ (contrairement à $m_K(z)$ qui ne dépend que des valeurs propres). L'auteur de [21] établit le résultat suivant :

Théorème 5. *Modulo quelques hypothèses techniques légères,*

$$q_K(z) - b_K(z) \xrightarrow{K \rightarrow \infty} 0 \quad \text{presque sûrement}$$

pour tout $z \in \mathbb{C}_+$, où $b_K(z)$ est la fonction déterministe

$$b_K(z) = \sum_{r=1}^{\bar{M}} \frac{\mathbf{a}^* \mathbf{\Pi}_r \mathbf{b}}{-z + \gamma_r(1 - c_K - c_K z p_K(z))} = \sum_{r=1}^{\bar{M}} \frac{\mathbf{a}^* \mathbf{\Pi}_r \mathbf{b}}{-z - \gamma_r z \tilde{p}_K(z)} \quad (1.17)$$

1.4.2 Résultats supplémentaires sur le support de π_K

Le théorème 3 nous dit en particulier que la mesure $\tilde{\pi}_K$ (et de manière équivalente π_K) possède partout une densité sauf éventuellement en zéro. Les commentaires qui suivent l'énoncé de ce théorème indiquent un moyen de déterminer le support de π . La structure de ce support appelle quelques remarques préliminaires. Intuitivement, si c est petit, $K \gg M$ et on peut s'attendre à ce que π soit « proche » de la mesure spectrale limite ν de $\mathbf{H}\mathbf{H}^*$. Cette intuition est confortée par le fait que si on considère le régime où M est fixe et $K \rightarrow \infty$, alors la suite π_K converge faiblement vers ν . Plus rigoureusement, en partant de l'équation (1.11), il est possible de prouver que quand $c \rightarrow 0$, la mesure π (indexée par c) converge faiblement vers la mesure limite ν de $\mathbf{H}\mathbf{H}^*$ [22, 26]. Ceci laisse penser que pour les petites valeurs de c , le support de π est constitué de \bar{M} intervalles disjoints qui supportent les \bar{M} Diracs dont est constituée la mesure ν .

Une caractérisation du nombre de ces intervalles (en fonction de c) et de leurs bornes est proposée dans [22]. Nous présentons en quelques mots cette approche en nous aidant des résultats du paragraphe 1.3.2 du présent document.

Rappelons que $\tilde{p}(z)$ est définie sur \mathbb{C}_+ comme l'unique solution de (1.10) dans \mathbb{C}_+ et peut être étendue à \mathbb{R}^* d'après le théorème 3. Il n'est pas difficile de prouver que la fonction

$$f_K(z) = -\frac{1}{\tilde{p}_K(z)}$$

est aussi définie sur $\mathbb{C}_+ \cup \mathbb{R}^*$. En nous limitant à $z = x \in \mathbb{R}^*$, l'équation (1.12) devient

$$x = f \left(1 + c \int \frac{t}{f-t} \nu(dt) \right) = f \left(1 + \frac{c}{M} \sum_{m=1}^{\bar{M}} I_m \frac{\gamma_m}{f - \gamma_m} \right) \quad (1.18)$$

En reprenant le commentaire qui suit l'énoncé du théorème 3 on constate que pour déterminer le support de π , il suffit de considérer l'équation

$$x(f) = f \left(1 + \frac{c}{M} \sum_{m=1}^{\bar{M}} I_m \frac{\gamma_m}{f - \gamma_m} \right),$$

de repérer les plages où cette fonction est croissante (notons que $-1/\tilde{p}$ est croissante) et de supprimer tous les points en ordonnée correspondant à ces plages. Le support de π sur \mathbb{R}^* est constitué des points restants sur l'axe des ordonnées. Les bornes des intervalles qui constituent le support de π coïncident donc avec les extrêma de $x(f)$. La dérivée de $x(f)$ est

$$x'(f) = 1 - c \int \left(\frac{t}{f-t} \right)^2 \nu(dt) = 1 - \frac{c}{M} \sum_{m=1}^{\bar{M}} I_m \left(\frac{\gamma_m}{f - \gamma_m} \right)^2$$

Dans [22], Mestre montre que cette dérivée s'annule en un nombre pair $2Q$ de points. Les ordonnées de ces points coïncident avec les bornes des intervalles qui constituent le support de π sur \mathbb{R}^* . Ces résultats sont rassemblés dans la proposition suivante :

Proposition 1. *L'équation*

$$\frac{1}{M} \sum_{m=1}^{\bar{M}} I_m \left(\frac{\gamma_m}{f - \gamma_m} \right)^2 = \frac{1}{c} \quad (1.19)$$

admet un nombre pair $2Q$ de solutions $f_1^- < f_1^+ \leq f_2^- < f_2^+ \leq \dots \leq f_Q^- < f_Q^+$ en comptant les multiplicités. Le support \mathcal{S}_K de π_K dans \mathbb{R}^* est l'union de Q intervalles compacts

$$\mathcal{S}_K = [x_1^-, x_1^+] \cup \dots \cup [x_Q^-, x_Q^+]$$

où $x_q^- = x(f_q^-)$ et $x_q^+ = x(f_q^+)$ pour $q = 1, \dots, Q$ sont données par (1.18) et satisfont $x_1^- < x_1^+ \leq x_2^- < x_2^+ \leq \dots \leq x_Q^- < x_Q^+$.

Cette proposition nous dit que le support de π dans \mathbb{R}^* est constitué de Q « clusters » où Q dépend de c . En inspectant les lieux des solutions de l'équation

(1.19), on constate que pour tout $m \in \{1, \dots, \bar{M}\}$, il existe un $q \in \{1, \dots, Q\}$ unique pour lequel $\gamma_m \in (f_q^-, f_q^+)$, comme le montre la figure 1.1. Nous pouvons donc associer une valeur propre donnée γ_m à un certain cluster $[x_q^-, x_q^+]$ du support de π . Cependant, cette correspondance n'est pas en général biunivoque : deux valeurs propres différentes de \mathbf{HH}^* peuvent correspondre à un même cluster. Cela dit, le nombre de clusters augmente en raison inverse de c , comme le montrent l'équation (1.19) et la figure 1.1, et il arrive un moment où chaque cluster est associé à une valeur propre unique γ_m . Ceci confirme l'intuition selon laquelle un grand nombre d'échantillons par dimension du signal reçu est nécessaire pour observer finement le spectre de \mathbf{HH}^* . Pour que le cluster associé à la valeur propre γ_m soit dissocié des autres clusters, il faut que

$$\frac{1}{c} > \xi_K(m) \quad (1.20)$$

où $\xi_K(m)$ est une quantité qui se calcule sans difficulté à partir de (1.19). L'expression de $\xi_K(m)$ se trouve dans [22, Eq. (15)].

Ces résultats qui concernent le support de π sont de nature déterministe. Venons-

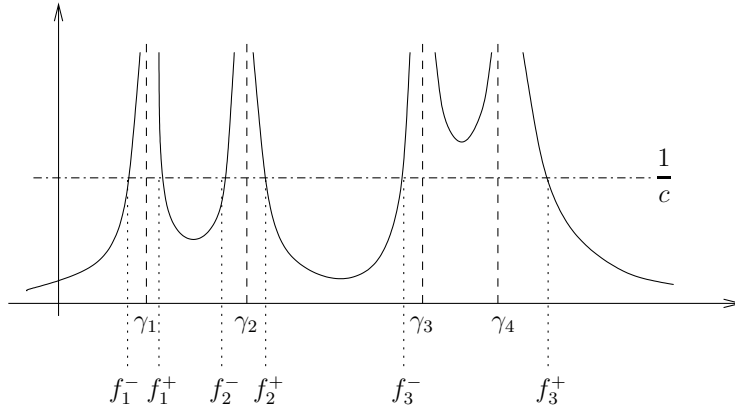


FIG. 1.1 – Représentation du membre de gauche de (1.19)

en aux valeurs propres (aléatoires) de $\Sigma\Sigma^*$. Dans [3], Bai et Silverstein raffinent le résultat de [2] et prouvent que pour tous grands K et avec probabilité un, la répartition des λ_i dans leurs clusters respectifs est exacte.

Nous abordons maintenant la question du changement de variable dans le calcul des intégrales (1.15) et (1.16).

1.4.3 Changement de variable et détermination du contour

Les propriétés suivantes de la fonction $f_K(z) = -1/\tilde{p}_K(z)$ se déduisent sans difficulté de la nature de $\tilde{p}(z)$. Notons que $\tilde{p}(z)$ s'étend à une fonction holomorphe sur $\mathbb{C} - \text{supp}(\tilde{\pi})$ et qui ne s'annule pas sur cet ensemble.

- La fonction $f(z)$ est holomorphe sur $\mathbb{C} - \mathcal{S}_K$ où $\mathcal{S}_K = [x_1^-, x_1^+] \cup \dots \cup [x_Q^-, x_Q^+]$.
- La partie imaginaire de f satisfait $\Im(f) > 0$ pour $x \in (x_1^-, x_1^+) \cup \dots \cup (x_Q^-, x_Q^+)$ et $\Im(f) = 0$ ailleurs sur \mathbb{R} .

D'autres propriétés sont prouvées dans [22] :

- La fonction $f(x)$ est continue et dérivable sur \mathbb{R} sauf aux points $\{x_1^-, x_1^+, \dots, x_Q^-, x_Q^+\}$.
Nous notons par $f'(x)$ sa dérivée sur $\mathbb{R} - \{x_1^-, x_1^+, \dots, x_Q^-, x_Q^+\}$ (cf aussi [26]).
- La partie réelle de $f(x)$, $x \in \mathbb{R}$ croît de $-\infty$ à $+\infty$ quand x croît de $-\infty$ à $+\infty$.
Par ailleurs, $f(x_q^-) = f_q^-$ et $f(x_q^+) = f_q^+$ où les (f_q^-, f_q^+) sont définis dans la proposition 1.

Pour résumer, $\Re(f(x))$ parcourt tout l'axe des réels quand x va de $-\infty$ à $+\infty$, $\Im(f(x)) > 0$ seulement si $\Re(f(x)) \in (f_1^-, f_1^+) \cup \dots \cup (f_Q^-, f_Q^+)$ et $\Im(f(x)) = 0$ ailleurs sur \mathbb{R} . Par conséquent, la fonction $f(x)$, concaténée avec sa conjuguée $f^*(x)$, décrit dans le plan complexe la courbe représentée sur la figure 1.2 quand x parcourt \mathbb{R} . On constate que si la condition de séparation (1.20) est satisfaite pour la valeur propre γ_m , alors la courbe engendrée par la concaténation de f et de f^* possède une portion qui entoure γ_m à l'exception de toute autre valeur propre. Dans toute la suite, nous

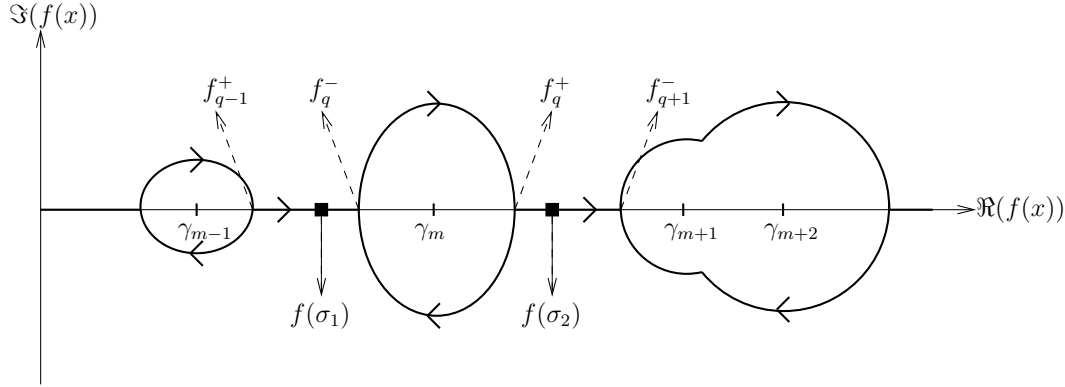


FIG. 1.2 – Lieux de $f(x)$ et $f^*(x)$ quand x parcourt \mathbb{R}

supposons que la condition (1.20) est satisfaite. Revenons aux intégrales (1.15) et (1.16). D'après ce qui précède, on constate qu'un contour \mathcal{C}_m admissible est décrit par les lieux de $f(x)$ où x va de σ_1 à σ_2 concaténés avec ceux de $f^*(x)$ où x revient de σ_2 vers σ_1 pour deux réels σ_1 et σ_2 bien choisis. Nous avons ainsi

$$\begin{aligned} \gamma_m &= \frac{M}{I_m} \frac{1}{2i\pi} \int_{\sigma_1}^{\sigma_2} \left(\frac{1}{M} \sum_{r=1}^{\bar{M}} K_r \frac{f(x)f'(x)}{\gamma_r - f(x)} \right) dx \\ &\quad - \frac{M}{I_m} \frac{1}{2i\pi} \int_{\sigma_1}^{\sigma_2} \left(\frac{1}{M} \sum_{r=1}^{\bar{M}} K_r \frac{f(x)f'(x)}{\gamma_r - f(x)} \right)^* dx \end{aligned} \quad (1.21)$$

et

$$\eta_m = \frac{1}{2i\pi} \int_{\sigma_1}^{\sigma_2} \left(\sum_{r=1}^{\bar{M}} \frac{\mathbf{a}^* \mathbf{\Pi}_r \mathbf{b}}{\gamma_r - f(x)} f'(x) \right) dx - \frac{1}{2i\pi} \int_{\sigma_1}^{\sigma_2} \left(\sum_{r=1}^{\bar{M}} \frac{\mathbf{a}^* \mathbf{\Pi}_r \mathbf{b}}{\gamma_r - f(x)} f'(x) \right)^* dx \quad (1.22)$$

Il s'agit maintenant d'estimer ces intégrales d'une manière consistante.

1.4.4 Passage dans le plan complexe

Regardons de plus près l'intégrande du premier terme de droite de (1.21). Porté en $z \in \mathbb{C}_+$, cet intégrande satisfait

$$\begin{aligned} \frac{1}{M} \sum_{r=1}^{\bar{M}} K_r \frac{f(z)f'(z)}{\gamma_r - f(z)} &= \frac{(\tilde{p}(z)^{-1})'}{\tilde{p}(z)} \int \frac{1}{t + \tilde{p}(z)^{-1}} \nu(dt) \\ &\stackrel{(a)}{=} -zp(z)(\tilde{p}(z)^{-1})' = zp(z) \frac{1 - c + cz^2 p'(z)}{(1 - c - czp(z))^2} \end{aligned}$$

où (a) est due à (1.9) et où $p'(z)$ est la dérivée complexe de $p(z)$ qui est, rappelons-le, holomorphe. De la même manière, nous avons

$$\sum_{r=1}^{\bar{M}} \frac{\mathbf{a}^* \mathbf{\Pi}_r \mathbf{b}}{\gamma_r - f(z)} f'(z) = b(z) \frac{1 - c + cz^2 p'(z)}{1 - c - czp(z)}$$

où $b(z)$ est donnée en (1.17). Les membres de droite de ces équations possèdent deux propriétés essentielles : ils ne dépendent pas explicitement des γ_r et il peuvent être estimés d'une manière consistante en tout point $z \in \mathbb{C}_+$ grâce aux théorèmes 2 et 5. L'idée est donc de réexprimer les intégrales (1.21) et (1.22) sous la forme d'intégrales de contour évaluées en dehors de l'axe des réels :

Proposition 2. *Si la condition de séparation (1.20) est satisfaite, les intégrales (1.21) et (1.22) possèdent les expressions suivantes :*

$$\begin{aligned} \gamma_m &= \frac{M}{I_m} \frac{1}{2i\pi} \oint_{\partial \mathbb{R}_y^-(m)} h_K(z) dz \\ \eta_m &= \frac{1}{2i\pi} \oint_{\partial \mathbb{R}_y^-(m)} g_K(z) dz \end{aligned} \tag{1.23}$$

où

$$h_K(z) = zp_K(z) \frac{1 - c_K + c_K z^2 p'_K(z)}{(1 - c_K - c_K zp_K(z))^2}, \quad g_K(z) = b_K(z) \frac{1 - c_K + c_K z^2 p'_K(z)}{1 - c_K - c_K zp_K(z)}$$

et où $\partial \mathbb{R}_y^-(m)$ est la frontière orientée négativement du rectangle

$$\mathbb{R}_y^-(m) = \{z \in \mathbb{C} : \sigma_1 \leq \Re(z) \leq \sigma_2, |\Im(z)| \leq y\} .$$

Cette proposition est prouvée dans [21] par une utilisation judicieuse du théorème de convergence dominée de Lebesgue. La principale difficulté provient du fait que $f'(x)$ n'est pas bornée aux voisinages des $x_q^{-,+}$.

L'étape suivante consiste à remplacer $h_K(z)$ et $b_K(z)$ par leurs estimées respectives.

1.4.5 Estimation consistante dans le cas centré : résultat final

En remplaçant $h_K(z)$ et $b_K(z)$ par leurs estimées

$$\hat{h}_K(z) = zm_K(z) \frac{1 - c_K + c_K z^2 m'_K(z)}{(1 - c_K - c_K z m_K(z))^2} \text{ et } \hat{g}_K(z) = q_K(z) \frac{1 - c_K + c_K z^2 m'_K(z)}{1 - c_K - c_K z m_K(z)} \quad (1.24)$$

dans les expressions des intégrales (1.23), on obtient des estimées consistantes de γ_m et de η_m :

Proposition 3. *Sous la condition de séparation (1.20),*

$$\begin{aligned} \gamma_m - \frac{M}{I_m} \frac{1}{2i\pi} \oint_{\partial R_y^-(m)} \hat{h}_K(z) dz &\xrightarrow[K \rightarrow \infty]{p.s.} 0 \\ \eta_m - \frac{1}{2i\pi} \oint_{\partial R_y^-(m)} \hat{g}_K(z) dz &\xrightarrow[K \rightarrow \infty]{p.s.} 0 . \end{aligned} \quad (1.25)$$

La preuve de cette proposition [21] fait appel aux résultats de séparation de Silverstein et Bai établis dans [2] et [3] : avec la probabilité un, les valeurs propres de $\Sigma \Sigma^*$ associées au cluster délimité par $\partial R_y^-(m)$ restent dans ce cluster pour tous K grands.

Il reste à donner des expressions explicites aux intégrales ci-dessus. Ces intégrales se traitent à l'aide de la formule des résidus. Par exemple, en revenant à (1.24) nous avons

$$\begin{aligned} &\frac{M}{I_m} \frac{1}{2i\pi} \oint_{\partial R_y^-(m)} \hat{h}_K(z) dz \\ &= -\frac{M}{I_m} \left(\sum_{\lambda_i \in R_y^-(m)} \text{Res}(\hat{h}_K(z), \lambda_i) + \sum_{\rho_i \in R_y^-(m)} \text{Res}(\hat{h}_K(z), \rho_i) \right) \end{aligned}$$

où nous rappelons que les λ_i sont les valeurs propres de $\hat{\mathbf{R}}_K = \Sigma_K \Sigma_K^*$, *i.e.*, les pôles de $p_K(z)$, et où les ρ_i sont les zéros des dénominateurs dans (1.24), qui s'avèrent être les solutions réelles de l'équation en ρ

$$\frac{1}{M} \sum_{i=1}^M \frac{\lambda_i}{\lambda_i - \rho} = \frac{1}{c} .$$

Grâce au principe de séparation établi dans [2, 3], avec la probabilité un, seules les valeurs propres $\{\lambda_i : i \in \mathcal{I}_m\}$ seront présentes dans la région d'intégration $R_y^-(m)$ pour K grand. Il existe un résultat similaire pour les ρ_k :

Lemme 1. *Sous la condition de séparation (1.20), avec la probabilité un, les $\{\rho_i : i \in \mathcal{I}_m\}$ se trouvent dans l'intervalle (σ_1, σ_2) et les $\{\rho_i : i \notin \mathcal{I}_m\}$ se trouvent en dehors de cet intervalle pour tous K grands.*

Ainsi, seuls les résidus associés aux λ_i et aux ρ_i pour $i \in \mathcal{I}_m$ entrent en jeu. Tous calculs faits, nous obtenons le théorème suivant :

Theorème 6. *Supposons la condition de séparation (1.20) satisfaite. Soit*

$$\hat{\gamma}_{m,K} = \frac{K}{I_m} \sum_{i \in \mathcal{I}_m} (\lambda_i - \rho_i), \quad \hat{\eta}_{m,K} = \sum_{i=1}^M \theta_m(i) \mathbf{a}^* \mathbf{v}_i \mathbf{v}_i^* \mathbf{b}$$

où

$$\begin{aligned} \theta_m(i) &= \begin{cases} -\phi_m(i), & i \notin \mathcal{I}_m \\ 1 + \psi_m(i), & i \in \mathcal{I}_m \end{cases} \\ \phi_m(i) &= \sum_{r \in \mathcal{I}_m} \left(\frac{\lambda_r}{\lambda_i - \lambda_r} - \frac{\rho_r}{\lambda_i - \rho_r} \right) \\ \psi_m(i) &= \sum_{r \notin \mathcal{I}_m} \left(\frac{\lambda_r}{\lambda_i - \lambda_r} - \frac{\rho_r}{\lambda_i - \rho_r} \right) \end{aligned}$$

et où $\rho_1 \leq \rho_2 \leq \dots \leq \rho_M$ sont les solutions réelles de l'équation en ρ

$$\frac{1}{M} \sum_{i=1}^M \frac{\lambda_i}{\lambda_i - \rho} = \frac{1}{c}.$$

en comptant les multiplicités. Quand $c > 1$, on utilise la convention $\rho_1 = \dots = \rho_{[M-K]+1} = 0$, et $\rho_{[M-K]+2}, \dots, \rho_M$ sont les solutions positives de cette équation. Dans le régime $K \rightarrow \infty$ et $M/K \rightarrow c > 0$,

$$\hat{\gamma}_{m,K} - \gamma_m \xrightarrow[K \rightarrow \infty]{p.s.} 0 \quad \text{et} \quad \hat{\eta}_{m,K} - \eta_m \xrightarrow[K \rightarrow \infty]{p.s.} 0.$$

Ces expressions peuvent être comparées à celles des estimateurs « classiques » (1.14) des mêmes quantités.

Avant de terminer ce paragraphe, notons qu'il est certainement possible d'évaluer les performances des estimateurs $\hat{\gamma}_{m,K}$ et $\hat{\eta}_{m,K}$ sous la forme d'un TLC. Pour réaliser cette étude, on pourrait partir des formes intégrales exprimées dans (1.25) et utiliser le TLC établi par Bai et Silverstein dans [4] (cf. §1.3.4).

1.5 Mesure spectrale du modèle Information plus bruit

1.5.1 Convergence de la mesure spectrale

Ce problème a été étudié dans [7, 14, 16] (entre autres contributions). Dans le cas où la mesure spectrale de $\mathbf{A}\mathbf{A}^*$ converge, le comportement asymptotique de la mesure spectrale de $\Sigma_K \Sigma_K^*$ converge lui aussi :

Theorème 7. *Pour $K = 1, 2, \dots$,*

– *Soit $M = M(K)$ une suite d'entiers telle que*

$$\frac{M(K)}{K} \xrightarrow[K \rightarrow \infty]{} c > 0.$$

- Soit $\mathbf{X}_K = [X_{mk}^{(K)}]$ une matrice aléatoire $M \times K$ dont les éléments $X_{mk}^{(K)} \in \mathbb{C}$ constituent un tableau $M \times K$ d'éléments iid, $\mathbb{E}X_{11}^{(K)} = 0$ et $\mathbb{E}|X_{11}^{(K)}|^2 = 1$.
- Soit \mathbf{A}_K une suite de matrices déterministes $M \times K$ telles que la mesure spectrale de $\mathbf{A}_K \mathbf{A}_K^*$ converge faiblement vers une mesure de probabilité ν quand $K \rightarrow \infty$.

Alors la mesure spectrale de $\Sigma_K \Sigma_K^*$ où $\Sigma_K = \mathbf{A}_K + \frac{1}{\sqrt{K}} \mathbf{X}_K$ converge presque sûrement en loi vers une mesure μ dont la transformée de Stieltjes $m(z)$ satisfait

$$m(z) = \int \frac{1}{\frac{t}{1 + cm(z)} - z(1 + cm(z)) + 1 - c} \nu(dt) \quad (1.26)$$

pour tout $z \in \mathbb{C}_+$.

Notons que ce résultat peut être généralisé au cas où la mesure spectrale de $\mathbf{A}_K \mathbf{A}_K$ ne converge pas forcément [16].

Le support de la mesure spectrale limite μ est étudié dans [6] :

Theorème 8. *Pour tout $x \in \mathbb{R}^*$, la limite $\lim_{z \in \mathbb{C}_+ \rightarrow x} m(z) \equiv m(x)$ existe. La fonction $m(x)$ est continue sur \mathbb{R}^* . La mesure de probabilité μ associée à $m(x)$ possède une densité continue f sur \mathbb{R}^* donnée par $f(x) = \frac{1}{\pi} \Im m(x)$. De plus, si $\Im m(x) > 0$ pour $x \neq 0$, alors $m(x)$ est une solution de (1.26) pour $z = x$, et la densité f est analytique dans un voisinage de x .*

Comme dans [26], les auteurs de [6] proposent une technique de détermination du support de μ à partir de l'inverse de $m(z)$.

Il s'avère possible de développer une approche d'estimation consistante pour le modèle Information plus bruit analogue à l'approche développée dans [21]. Cette approche sera décrite dans les rapports des WP2 et WP3. Afin de parfaire cette étude, il nous sera nécessaire d'établir des résultats relatifs aux valeurs propres extrêmes analogues à ceux de [2] et [3]. Par ailleurs, l'étude des performances asymptotiques des estimateurs qui seront proposés réclame un TLC (comme dans [4]). Ces travaux sont en cours.

Chapitre 2

Estimation de matrices de covariance structurées

Il y a près de deux décennies, un certain nombre de techniques qui visent à améliorer l'estimation de matrices de covariances associées à des séries temporelles ont été introduites. Les premières références importantes qui traitèrent de ce sujet sont [10, 19, 8, 29], travaux qui furent suivis par [28, 11, 9].

L'idée est que les matrices de covariance théoriques possèdent souvent des propriétés structurelles qui ne se retrouvent plus dans les estimées traditionnelles de ces matrices de covariance (telles les matrices de covariance empiriques). Par exemple, dans un contexte basique de séries temporelles, considérons un signal stationnaire qui consiste en un nombre fini de sinusôides auxquelles se rajoute un bruit blanc. La matrice de covariance possède deux types de structure : la structure Toeplitz due à la stationnarité, et un certain profil de valeurs propres (la valeur propre due à la variance du bruit ayant une certaine multiplicité). Ces deux propriétés structurelles disparaissent dans une matrice de covariance empirique correspondante. On peut imaginer qu'une estimée améliorée de la matrice de covariance qui possède ces propriétés structurelles puisse assurer une meilleure extraction des paramètres du signal. De là l'idée d'améliorer l'estimée de la matrice de covariance de façon à renforcer ces propriétés structurelles.

Il est difficile de renforcer simultanément toutes les propriétés structurelles. De là, un procédé itératif a été proposé selon lequel les propriétés structurelles sont partitionnées et chaque sous-ensemble de propriétés est renforcé tour à tour. Chaque sous-ensemble de propriétés correspond typiquement à un ensemble convexe, de manière à ce que le renforcement de la structure correspondante se ramène à une projection sur cet ensemble convexe. Il en résulte que le processus itératif converge vers l'intersection de ces ensembles convexes. Cette intersection satisfait toutes les contraintes structurelles et se trouve être le plus proche de la matrice de covariance originelle.

Plus récemment, les versions multi canal de ces approches ont suscité de l'intérêt. En effet, dans bon nombre de problèmes, nous devons considérer des matrices de covariance de processus vectoriels. Ces matrices correspondent typiquement aux signaux d'un nombre limité de sources qui ont été filtrés par un canal matriciel avant d'atteindre un certain nombre de capteurs en présence de bruit. Certaines applications

se rencontrent dans le domaine du traitement audio où les sources sont des haut-parleurs et les capteurs des microphones, ou dans le domaine des communications sans fil où les sources sont des antennes d'émission et les capteurs sont des antennes de réception. Si ces signaux sont stationnaires, leur spectre matriciel peut avoir un rang faible à toutes les fréquences si le nombre de sources est inférieur au nombre de capteurs. D'un autre côté, les corrélations dans le domaine temporel possèdent un étalement temporel fini si les filtres qui représentent les canaux sont à mémoire finie, ce qui est typiquement le cas. Les procédures d'amélioration des estimées des covariances dans le cas multi canal sont en cours d'étude dans le cadre du WP4 de SESAME.

La caractérisation des performances des covariances améliorées est un sujet difficile. Il est souhaitable de simplifier ce problème en considérant un régime asymptotique approprié où la taille de la matrice de covariance (ou le nombre de points dans le domaine fréquentiel) tend vers l'infini et une autre dimension du problème croît proportionnellement de manière à ce que le comportement limite demeure pertinent pour les problèmes de dimension finie, où toutes les dimensions se trouvent dans des rapports finis. Parmi les autres dimensions du problème, nous pouvons compter le rang de la matrice de covariance ou le support temporel de la fonction de corrélation. De telles analyses asymptotiques font partie des tâches assignées au WP4.

Références bibliographiques

- [1] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields*, 134(2) :283–338, 2006.
- [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] Z.D. Bai and J.W. Silverstein. Exact separation of eigenvalues of large dimensional sample covariance matrices. *Annals of Probability*, 27(3) :1536–1555, 1999.
- [4] Z.D. Bai and J.W. Silverstein. CLT of linear spectral statistics of large dimensional sample covariance matrices. *Annals of Probability*, 32(1A) :553–605, 2004.
- [5] A. Boutet de Monvel and A. Khorunzhy. Limit theorems for random matrices. *Markov Process. Related Fields*, 4(2) :175–197, 1998.
- [6] R. Brent Dozier and J.W. Silverstein. Analysis of the Limiting Spectral Distribution of Large Dimensional Information-Plus-Noise Type Matrices. *Journal of Multivariate Analysis*, 2007.
- [7] R. Brent Dozier and J.W. Silverstein. On the Empirical Distribution of Eigenvalues of Large Dimensional Information-Plus-Noise Type Matrices. *Journal of Multivariate Analysis*, (98) :678–694, 2007.
- [8] J. A. Cadzow. Signal enhancement - A composite property mapping algorithm. *IEEE Trans. Acoust., Speech and Signal Processing*, 36(1), 1988.
- [9] J.A. Cadzow and D.M. Wilkes. Enhanced rational signal modeling. *Signal Processing*, 25 :171–188, 1991.
- [10] M.P. Ekstrom. An iterative-improvement approach to the numerical solution of vector Toeplitz systems. *IEEE Trans. Comput.*, 23 :320–325, 1975.
- [11] D.R. Fuhrmann and M.I. Miller. On the existence of positive-definite maximum-likelihood estimate of structured covariance matrices. *IEEE Trans. on Information Theory*, 34(4) :722–729, 1988.
- [12] V.L. Girko. *Theory of Random Determinants*, volume 45 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990.
- [13] V.L. Girko. *An Introduction to Statistical Analysis of Random Arrays*. VSP, The Netherlands, 1998.

- [14] V.L. Girko. *Theory of Stochastic Canonical Equations*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001. Two volumes.
- [15] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur. A new approach for mutual information analysis of large dimensional multi-antenna channels. *IEEE Trans. on Information Theory*, 54(9) :3987–4004, Sep. 2008.
- [16] W. Hachem, Ph. Loubaton, and J. Najim. Deterministic Equivalents for Certain Functionals of Large Random Matrices. *Annals of Applied Probability*, 17(3), 2007.
- [17] W. Hachem, Ph. Loubaton, and J. Najim. A CLT for information-theoretic statistics of Gram random matrices with a given variance profile. *Annals of Applied Probability*, 18(6) :2071–2130, 2008.
- [18] F. Hiai and D. Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.
- [19] S.Y. Kung, C.K. Lo, and R. Foka. A Toeplitz approximation approach to coherent source direction finding. In *Proc. ICASSP*, Tokyo, Japan, 1986.
- [20] V.A. Marčenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Math. USSR Sbornik*, 1(4) :457–483, 1967.
- [21] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *IEEE Trans. on Information Theory*, 54(11) :5113–5129, November 2008.
- [22] X. Mestre. On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices. *IEEE Trans. on Signal Processing*, 56(11) :5353–5368, November 2008.
- [23] Projet ANR SESAME. Fiche B : « Description technique du projet » . Réponse à l'appel à projets ANR « Masses de Données et COonnaissances » (ANR-07-MDCO), 2007.
- [24] J.W Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.*, 55(2) :331–339, 1995.
- [25] J.W. Silverstein and Z.D. Bai. On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2) :175–192, 1995.
- [26] J.W Silverstein and Sang-II Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *J. Multivariate Anal.*, 54(2) :295–309, 1995.
- [27] D. Voiculescu. Limit laws for random matrices and free products. *Inventiones Mathematicae*, 104 :201–220, 1991.
- [28] D.M. Wilkes and M.H. Hayes. Block Toeplitz approximation. *Signal Processing*, 15 :303–313, 1988.
- [29] D.M. Wilkes and M.H. Hayes. Iterated Toeplitz approximation of covariance matrices. In *Proc. ICASSP*, New York, USA, 1988.

- [30] Y. Q. Yin. Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.*, 20(1) :50–68, 1986.