Sur les fluctuations de l'information mutuelle des canaux MIMO de grandes dimensions

Walid HACHEM[1], Philippe LOUBATON[2], Jamal NAJIM[1]

[1]CNRS, Télécom Paris (UMR 5141)
46, rue Barrault, 75634 Paris Cedex 13, France

[2]Université de Marne-La-Vallée
5, Blvd Descartes, Champs-Sur-Marne, 77454 Marne-La-Vallée, France

walid.hachem@enst.fr, loubaton@univ-mlv.fr, jamal.najim@enst.fr

Résumé – Dans cet article, la théorie des grandes matrices aléatoires est utilisée pour étudier l'information mutuelle de Shannon d'une classe générale de canaux radio MIMO (“ Multiple Input Multiple Output ”) à gains aléatoires corrélés. Il existe dans la littérature une approximation de cette information mutuelle dans le régime asymptotique où le nombre d'antennes d'émission et le nombre d'antennes de réception tendent vers l'infini au même rythme. Cette contribution s'attache à étudier les fluctuations de l'information mutuelle autour de cette approximation déterministe sous la forme d'un Théorème de la Limite Centrale. Ce TLC peut en particulier servir d'approximation de la probabilité de coupure (“ Outage Probability ”), critère de performance fondamental dans le cas où les évanouissements du canal sont lents.

Abstract – In this article, large random matrix theory is used to study Shannon's mutual information of a general class of MIMO (“Multiple Input Multiple Output”) radio channels with random correlated gains. In the literature there exists an approximation of this mutual information in the asymptotic regime where the number of transmitting antennas and the number of receiving antennas grow toward infinity at the same pace. This contribution is devoted to the study of the mutual information's fluctuations around this deterministic approximation in the form of a Central Limit Theorem. In particular, this CLT can be used to approximate in a simple manner the Outage Probability, which is a fundamental performance criterion in the case where channel fades are slow.

1 Introduction

Dans un contexte de transmission sans fil où un grand nombre de réflecteurs et de diffuseurs se trouvent sur le chemin des ondes radio, l'utilisation d'antennes multiples en émission et en réception accroît notablement l'efficacité spectrale. Dans ce contexte de transmission, il est pertinent d'assimiler les éléments de la matrice qui représente le canal MIMO (“ Multiple Input Multiple Output ”) à des variables aléatoires. En supposant que ces éléments sont des variables aléatoires gaussiennes indépendantes et identiquement distribuées (iid), Telatar a en effet réalisé vers le milieu des années quatre-vingt-dix que la capacité de Shannon du canal croît selon $\min(r,t)$ où r est le nombre d'antennes en réception et t le nombre d'antennes en émission, la puissance totale à l'émission étant fixée [12]. Soit $\mathbf{H}$ la matrice $r \times t$ qui représente le canal MIMO étudié dans [12] et soit $I(\rho) = \mathbb{E}\mathcal{I}(\rho)$ la capacité de Shannon par antenne de réception de ce canal où $\mathcal{I}(\rho)$ désigne la variable aléatoire $\mathcal{I}(\rho) = \frac{1}{r} \log\det\left(\frac{1}{\rho t}\mathbf{H}\mathbf{H}^H + \mathbf{I}_r\right)$ et ρ est la variance d'un bruit blanc gaussien additif. Un résultat classique de la théorie des matrices aléatoires affirme que si les éléments de $\mathbf{H}$ sont iid, la distribution empirique des valeurs propres de $\mathbf{H}\mathbf{H}^H$ converge faiblement vers une loi déterministe quand $t \to \infty$ et r/t tend vers une constante $c > 0$ (tout au long de l'article, “ $t \to \infty$ ” dénotera ce régime asymptotique) [9]. En se basant sur le fait que la capacité par antenne de réception $I(\rho)$ est l'espérance d'une fonction log par rapport à cette loi empirique, il est alors possible d'établir la convergence de $I(\rho)$ vers une constante. Ce résultat précise l'assertion selon laquelle la capacité croît comme $\min(r,t)$ quand $t \to \infty$.
Les résultats de [12] ont ensuite été généralisés pour traiter divers modèles statistiques de canaux décrivant les corrélations entre les éléments de $\mathbf{H}$ que l'on constate dans la pratique. Sans être exhaustifs, citons parmi ces généralisations [6, 7, 4, 10, 13]. Ces travaux exhibent une approximation $V(\rho)$ de l'information mutuelle dans le sens où $I(\rho) - V(\rho) \to 0$ quand $t \to \infty$. Notons tout de suite que $V(\rho)$ ne possède une expression analytique que dans quelques cas “ académiques ” dont celui traité par Telatar. En général, $V(\rho)$ est liée à la solution d'un système d'équations implicites comme nous le verrons plus loin.
Une question essentielle dans l'étude de l'information mutuelle par l'intermédiaire de la théorie des matrices aléatoires consiste à caractériser les fluctuations de la variable aléatoire $\mathcal{I}(\rho) - V(\rho)$. Cette étude prendra la forme d'un Théorème de la Limite Centrale (TLC), *i.e.*, on cherchera à établir le fait que $r^\alpha(\mathcal{I}(\rho) - V(\rho)) - \mathcal{N}(b_t, \Theta_t) \to 0$ en loi quand $t \to \infty$ où $\mathcal{N}(b_t, \Theta_t)$ désigne une variable aléatoire gaussienne de moyenne b_t et de variance Θ_t et le coefficient α caractérise la vitesse de convergence. En plus de

l'importance évidente de ce résultat pour juger la validité de l'approximation $V(\rho)$ pour un nombre fini d'antennes, il fournit une approximation de la probabilité dite de coupure (" Outage Probability ") $\mathbb{P}[r\mathcal{I}(\rho) \leq R]$ où R est un débit donné. Rappelons que dans un contexte de faible mobilité où les évanouissements du canal radio sont lents, le critère de performance pertinent est la probabilité de coupure et non l'information mutuelle $I = \mathbb{E}\mathcal{I}$.

Dans cet article, nous considérons un modèle selon lequel la matrice $\mathbf{H}$ se ramène à une matrice $\mathbf{Y}$ à éléments gaussiens indépendants centrés et de variances quelconques (non forcément égales). Ce modèle englobe la quasi totalité des modèles MIMO *centrés* étudiés dans la littérature des communications radio. Les résultats mathématiques existants (citons [1, 2]) ne répondent au problème des fluctuations de $\mathcal{I}(\rho) - V(\rho)$ que dans quelques cas particuliers du modèle adopté dans cet article. En ce sens, cet article apporte une nouvelle contribution à la théorie des matrices aléatoires.

Le paragraphe 2 est consacré à la présentation du modèle statistique de $\mathbf{H}$ et au rappel de l'expression de l'approximation $V(\rho)$ présente dans la littérature. Notre TLC fera l'objet du paragraphe 3. En particulier, la vitesse de convergence est $\alpha = 1$, le biais b_t est nul et une expression compacte de la variance Θ_t est donnée. Le principe de la preuve de ce TLC est présenté sommairement dans le paragraphe 4. La preuve détaillée se trouve dans [8]. Notons que dans ce dernier article, nous ne nous limitons pas au cas où les élements de $\mathbf{Y}$ sont gaussiens.

2 Modèle de canal et expression de $V(\rho)$

On considère un système de communication muni de t antennes d'émission et de r antennes de réception, et on désigne par $\mathbf{H}$ la matrice $r \times t$ représentant les gains complexes entre les différents émetteurs et récepteurs. En un instant donné, le vecteur $\mathbf{y}$ de dimension r reçu est décrit par l'équation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$$

où $\mathbf{z}$ est un vecteur aléatoire gaussien complexe de variance ρ représentant le bruit thermique, et où $\mathbf{x}$ représente le vecteur de dimension t transmis par l'émetteur. En supposant que $\mathbf{x}$ soit un vecteur gaussien centré de matrice de covariance $\mathbb{E}\mathbf{x}\mathbf{x}^H = \frac{1}{t}\mathbf{I}_t$, l'information mutuelle de Shannon du canal par antenne de réception est

$$I(\rho) = \frac{1}{r}\mathbb{E}\log\det\left(\frac{1}{\rho t}\mathbf{H}\mathbf{H}^H + \mathbf{I}_r\right) .$$

Il convient d'évaluer $I(\rho)$ pour différents modèles statistiques du canal $\mathbf{H}$. Ces modèles rendent compte des corrélations entre les gains du canal dues le plus souvent à la proximité des antennes au sein des réseaux d'émission et/ou de réception. Le modèle général de corrélation considéré dans ce travail est le suivant [11] : la matrice du canal s'écrit sous la forme $\mathbf{H} = \mathbf{F}_r\mathbf{Y}\mathbf{F}_t$ où

- Pour un entier donné n, la matrice $\mathbf{F}_n$ est la matrice de Fourier $n \times n$, *i.e.*, la matrice dont l'élément (k, l) est $\exp(2\imath\pi(k-1)(\ell-1)/n)/\sqrt{n}$ pour $k, \ell = 1, \ldots, n$,
- La matrice $\mathbf{Y}$ est une matrice aléatoire dont l'élément (k, ℓ) s'écrit $Y_{k\ell} = \sigma_{k\ell}X_{k\ell}$ où les variables aléatoires $X_{k\ell}$ sont indépendantes et de loi gaussienne complexe circulaire standard $\mathcal{CN}(0, 1)$ et les $\sigma_{k\ell}$ sont des nombres réels. Comme $\mathbb{E}|Y_{k\ell}|^2 = \sigma_{k\ell}^2$, le tableau $(\sigma_{k\ell}^2)_{k,\ell=1}^{r,t}$ est appelé un champ de variance.

En d'autres termes, $\mathbf{H}$ est la transformée de Fourier à deux dimensions d'une matrice à éléments indépendants sujette à un champ de variance. Selon ce modèle, les colonnes de la matrice $\mathbf{F}_r$ (resp. de $\mathbf{F}_t = \mathbf{F}_t^T$) jouent en quelque sorte le rôle de vecteurs de direction d'arrivée (resp. de départ) élémentaires. L'élément $Y_{k\ell}$ de $\mathbf{Y}$ matérialise le " trajet " vu sous les angles d'arrivée et de départ représentés par la colonne k de $\mathbf{F}_r$ et par la colonne ℓ de $\mathbf{F}_t$ respectivement. Comme les matrices de Fourier sont des matrices unitaires, *i.e.*, $\mathbf{F}_n\mathbf{F}_n^H = \mathbf{I}_n$, il est clair que $\log\det\left(\frac{1}{\rho t}\mathbf{H}\mathbf{H}^H + \mathbf{I}_r\right) = \log\det\left(\frac{1}{\rho t}\mathbf{Y}\mathbf{Y}^H + \mathbf{I}_r\right)$. Par conséquent, le problème revient à évaluer $I(\rho) = \mathbb{E}\mathcal{I}(\rho)$ où

$$\mathcal{I}(\rho) = \frac{1}{r}\log\det\left(\frac{1}{\rho t}\mathbf{Y}\mathbf{Y}^H + \mathbf{I}_r\right) . \tag{1}$$

En général, le calcul direct de $I(\rho)$ est extrêmement difficile. Afin de contourner cette difficulté, il est pertinent de chercher une approximation de $I(\rho)$. Grâce à la théorie des matrices aléatoires de grandes dimensions, il se révèle possible de mettre en évidence une quantité déterministe d'expression relativement simple $V(\rho)$ telle que $I(\rho) - V(\rho) \to 0$ quand $t \to \infty$ et $r/t \to c > 0$.

Dans toute la suite, nous aurons besoin de l'hypothèse technique suivante :

A1 : Il existe un réel $\sigma_{\max}$ tel que

$$\sup_{t\geq 1}\max_{\substack{1\leq k\leq r\\ 1\leq \ell\leq t}} \sigma_{k\ell}^2 < \sigma_{\max}^2 .$$

Par ailleurs, nous dirons qu'une fonction complexe $g(z)$ appartient à la classe $\mathcal{S}$ si $g(z)$ est analytique dans le demi plan $\mathbb{C}_+ = \{z \in \mathbb{C}; \mathrm{im}(z) > 0\}$, si $g(z) \in \mathbb{C}_+$ pour tout $z \in \mathbb{C}_+$ et si $\mathrm{im}(z)|g(z)|$ est bornée dans $\mathbb{C}_+$. Afin de donner l'expression de $V(\rho)$, nous avons besoin du résultat suivant [5, 7] :

Proposition 1 *Supposons* **A1** *vraie. Le système des r équations fonctionnelles*

$$g_k(z) = \frac{1}{-z + \frac{1}{t}\sum_{j=1}^{t}\frac{\sigma_{kj}^2}{1 + \frac{1}{t}\sum_{\ell=1}^{r}\sigma_{\ell j}^2 g_\ell(z)}}, \quad k = 1, \ldots, r \tag{2}$$

admet une solution unique $(g_1(z), \cdots, g_r(z))$ pour $g_k(z) \in \mathcal{S}$.

Les fonctions $g_k(z)$ ainsi définies se prolongent analytiquement sur l'ensemble $\mathbb{C} - [0, \infty)$.

L'approximation déterministe est donnée par le théorème suivant :

Théorème 1 *Supposons* **A1**. *Soit*

$$V(\rho) = -\frac{1}{r}\sum_{i=1}^{r}\log(\rho g_i(-\rho)) + \frac{1}{r}\sum_{j=1}^{t}\log\left(1+\frac{1}{t}\sum_{\ell=1}^{r}\sigma_{\ell j}^2 g_\ell(-\rho)\right) - \frac{1}{rt}\sum_{i=1:r,j=1:t}\frac{\sigma_{ij}^2 g_i(-\rho)}{1+\frac{1}{t}\sum_{\ell=1}^{r}\sigma_{\ell j}^2 g_\ell(-\rho)}$$

où les fonctions $g_i(z)$ sont définies par la proposition 1. Alors $I(\rho) - V(\rho) \to_{t\to\infty} 0$ pour tout $\rho > 0$.

Ce résultat se trouve dans [13] et dans [7]. Il se base sur une description du comportement asymptotique de la distribution empirique des valeurs propres de la matrice $\mathbf{YY}^H$ réalisée par le biais des fonctions $g_i(z)$ Le premier résultat relatif à ce comportement asymptotique remonte sans doute à [5].

3 Fluctuations de $\mathcal{I}(\rho) - V(\rho)$

La quantité $rI(\rho) = r\mathbb{E}\mathcal{I}(\rho)$ appelée parfois " information mutuelle ergodique " est la borne supérieure du débit que le canal $\mathbf{H}$ peut théoriquement assurer dans le cas où la durée du mot de code est largement supérieure au temps de cohérence de ce canal. Dans le cas contraire où le canal varie peu le long d'un mot de code, l'indice de performance pertinent n'est plus $rI(\rho)$ mais plutôt la probabilité de coupure $\mathbb{P}\left[r\mathcal{I}(\rho) < R\right]$ où R est le débit visé. Jusque là, la théorie des matrices aléatoires nous a seulement fourni une approximation $V(\rho)$ de l'information mutuelle ergodique. Afin d'évaluer la pertinence de l'approximation $V(\rho)$ et en même temps d'obtenir une approximation de la probabilité de coupure, on cherche à prouver que dans le régime asymptotique, la quantité $r(\mathcal{I}(\rho)-V(\rho))$ se comporte comme une variable aléatoire gaussienne.

Notre but est maintenant d'étudier le comportement asymptotique de la suite de variables aléatoires $X_t = r(\mathcal{I}(\rho) - V(\rho))$ que nous mettrons sous la forme $X_t = Z_t + b_t$ où $Z_t = r(\mathcal{I}(\rho) - I(\rho))$ est une variable aléatoire qui rend compte des fluctuations de $r\mathcal{I}(\rho)$ autour de sa moyenne et $b_t = r(I(\rho) - V(\rho))$ est un terme déterministe qui représente un biais. La contribution de ce travail est la suivante :

Théorème 2 *Soit la matrice $\mathbf{Y} = [Y_{k\ell}]$ de dimensions $r\times t$ où $Y_{k\ell} = \sigma_{k\ell}X_{k\ell}$, les variables aléatoires $\{X_{k\ell}\}_{k,\ell=1}^{r,t}$ étant indépendantes de loi circulaire gaussienne $\mathcal{CN}(0,1)$. Soit $\mathbf{G}$ la matrice diagonale $r\times r$*

$$\mathbf{G} = \mathrm{diag}\left([g_1(-\rho),\ldots,g_r(-\rho)]\right)$$

où les fonctions g_k sont définies par la proposition (1) et $\mathbf{C}_l$ les matrices diagonales $r\times r$

$$\mathbf{C}_\ell = \mathrm{diag}\left([\sigma_{1\ell}^2,\ldots,\sigma_{r\ell}^2]\right) \text{ pour } \ell = 1,\ldots,t\ .$$

Supposons **A1** *et*

$$\mathbf{A2}:\quad \max\left(\liminf_{t\geq 1}\min_{1\leq\ell\leq t}\frac{1}{t}\sum_{k=1}^{r}\sigma_{k\ell}^2\ ,\ \liminf_{t\geq 1}\min_{1\leq k\leq r}\frac{1}{n}\sum_{\ell=1}^{t}\sigma_{k\ell}^2\right) > 0\ .$$

Soit $\mathbf{A}$ la matrice de taille $t\times t$ définie par

$$\mathbf{A} = \left[\frac{1}{t}\frac{\frac{1}{t}\mathrm{tr}(\mathbf{C}_\ell\mathbf{C}_m\mathbf{G}^2)}{\left(1+\frac{1}{t}\mathrm{tr}\left(\mathbf{C}_\ell\mathbf{G}\right)\right)^2}\right]_{\ell,m=1}^{t}\ .$$

Alors les résultats suivants sont vrais :

- *Le réel $\Theta_t = -\log\det\left(\mathbf{I}_t - \mathbf{A}\right)$ est bien défini et*
$$0 < \liminf_t \Theta_t \leq \limsup_t \Theta_t < \infty\ .$$
- *Soit la suite de variables aléatoires $Z_t = r(\mathcal{I}(\rho) - \mathbb{E}\mathcal{I}(\rho))$ où $\mathcal{I}(\rho)$ est donnée par (1). Alors*
$$\frac{Z_t}{\sqrt{\Theta_t}} \xrightarrow[t\to\infty,r/t\to c>0]{} \mathcal{N}(0,1) \quad \text{en loi}\ .$$
- *Le biais $b_t = r(\mathbb{E}\mathcal{I}(\rho) - V(\rho))$ où $V(\rho)$ est donnée dans l'énoncé du théorème 1 satisfait*
$$b_t \xrightarrow[t\to\infty,r/t\to c>0]{} 0\ .$$

Ce théorème nous dit en résumé que $r(\mathcal{I}(\rho) - V(\rho))$ peut être approchée dans le régime des grandes dimensions par une variable aléatoire gaussienne $\mathcal{N}(0,\Theta_t)$.

Remarques

- L'énoncé du théorème 2 n'impose aucune restriction sur la structure du champ de variance (σ_{kl}^2) (en dehors des hypothèses **A1** et **A2**). Dans la littérature, des résultats sur les fluctuations de $r(\mathcal{I}(\rho) - V(\rho))$ existent dans le cas dit de Kronecker (ou séparable) où $\sigma_{k\ell}^2$ admet une factorisation de la forme $\sigma_{k\ell}^2 = \sigma_{\mathrm{R},k}^2\sigma_{\mathrm{T},\ell}^2$. Citons à ce propos [10] et [6]. Dans le cas séparable, le système de r équations (2) se ramène en un système à deux équations et la variance est de la forme $\Theta_t = -\log(1-\xi(\rho))$ où $\xi(\rho)$ est une quantité scalaire facile à calculer.
- Le théorème 2 admet une généralisation pour le cas où les éléments de $\mathbf{Y}$ ne sont pas nécessairement gaussiens [8]. Il serait utile de noter que dans ce cas on a toujours $\alpha = 1$ mais on constate l'apparition d'un biais ($b_t \neq 0$) et d'un facteur correctif pour la variance Θ_t, tous deux proportionnels au cumulant d'ordre quatre des $X_{k\ell}$. Nous ne nous attardons pas sur ce dernier point dans la présente contribution au vu de son importance limitée dans le contexte des communications sans fil considéré ici.

4 Principe de la preuve du th.2

Nous terminons par quelques indications sur la preuve du théorème 2. Nous nous intéressons à la variable aléatoire $Z_t = r(\mathcal{I}(\rho) - \mathbb{E}\mathcal{I}(\rho))$. Désignons par $\mathbf{y}_\ell$ la colonne ℓ de la matrice $\mathbf{Y}$ et notons par $\mathbb{E}_\ell$ l'espérance conditionnelle $\mathbb{E}_\ell[.] = \mathbb{E}[.\|\mathbf{y}_\ell,\ldots,\mathbf{y}_t]$. Nous pouvons écrire

$$Z_t = \sum_{\ell=1}^{t}\left(\mathbb{E}_\ell - \mathbb{E}_{\ell+1}\right)\log\det\left(\frac{1}{\rho t}\mathbf{YY}^H + \mathbf{I}_r\right) \stackrel{\mathrm{def}}{=} \sum_{\ell=1}^{t}W_\ell$$

où nous posons $\mathbb{E}_{t+1} = \mathbb{E}$. Par construction, la suite des variables aléatoires $(W_t, W_{t-1},\ldots,W_1)$ est une suite d'incréments de martingale par rapport à la suite de tribus

$\sigma(\mathbf{y}_t), \ldots, \sigma(\mathbf{y}_t, \ldots, \mathbf{y}_1)$ (*cf.* [3]). Le comportement asymptotique de la somme $Z_t = \sum_\ell W_\ell$ peut alors être caractérisé par le biais du TLC pour les martingales [3, Ch. 35]. Nous commençons par remanier les expressions des W_ℓ. Désignons par $\mathbf{Y}_\ell$ la matrice $\mathbf{Y}$ privée de la colonne $\mathbf{y}_\ell$. A l'évidence, nous avons

$$(\mathbb{E}_\ell - \mathbb{E}_{\ell+1}) \left[\log\det\left(\frac{1}{\rho t}\mathbf{Y}_\ell^H \mathbf{Y}_\ell + \mathbf{I}_{t-1}\right)\right] = 0$$

Par conséquent, comme $\det(\mathbf{B}\mathbf{B}^H + \mathbf{I}) = \det(\mathbf{B}^H\mathbf{B} + \mathbf{I})$, W_ℓ s'écrit

$$\begin{aligned} W_\ell &= (\mathbb{E}_\ell - \mathbb{E}_{\ell+1}) \log\left(\frac{\det\left(\frac{1}{\rho t}\mathbf{Y}^H\mathbf{Y} + \mathbf{I}_t\right)}{\det\left(\frac{1}{\rho t}\mathbf{Y}_\ell^H\mathbf{Y}_\ell + \mathbf{I}_{t-1}\right)}\right) \\ &\stackrel{\text{def}}{=} (\mathbb{E}_\ell - \mathbb{E}_{\ell+1}) \log\left(\frac{\det \mathbf{\Xi}}{\det \mathbf{\Xi}_\ell}\right) . \end{aligned}$$

En rappelant que $\det\begin{bmatrix} a & \mathbf{b}^H \\ \mathbf{b} & \mathbf{B} \end{bmatrix} = (a - \mathbf{b}^H\mathbf{B}^{-1}\mathbf{b})\det\mathbf{B}$, nous avons

$$\begin{aligned} \det\mathbf{\Xi} = (\det\mathbf{\Xi}_\ell)\Bigg(&\frac{\|\mathbf{y}_\ell\|^2}{\rho t} + 1 \\ &- \frac{\mathbf{y}_\ell^H\mathbf{Y}_\ell}{\rho t}\left(\frac{1}{\rho t}\mathbf{Y}_\ell^H\mathbf{Y}_\ell + \mathbf{I}\right)^{-1}\frac{\mathbf{Y}_\ell^H\mathbf{y}_\ell}{\rho t}\Bigg) . \end{aligned}$$

En utilisant la relation $\mathbf{I} - \mathbf{B}\left(\mathbf{B}^H\mathbf{B} + \mathbf{I}\right)^{-1}\mathbf{B}^H = (\mathbf{B}\mathbf{B}^H + \mathbf{I})^{-1}$, nous obtenons

$$W_\ell = (\mathbb{E}_\ell - \mathbb{E}_{\ell+1}) \log\left(1 + \frac{1}{t}\mathbf{y}_\ell^H\mathbf{Q}_\ell\mathbf{y}_\ell\right) \qquad (3)$$

où $\mathbf{Q}_\ell$ est la matrice résolvante $\mathbf{Q}_\ell = \left(\frac{1}{t}\mathbf{Y}_\ell\mathbf{Y}_\ell^H + \rho\mathbf{I}_r\right)^{-1}$. Un résultat fondamental qui remonte à [9] (voir aussi [2, 5, 7, 8]) établit que si $\mathbf{x}$ est un vecteur aléatoire $r \times 1$ dont les éléments sont centrés iid de variance 1 et si $\mathbf{B}$ est une matrice hermitienne $r \times r$ indépendante de $\mathbf{x}$, alors $\frac{1}{t}(\mathbf{x}^H\mathbf{B}\mathbf{x} - \mathrm{tr}(\mathbf{B})) \to 0$ quand $t \to \infty$. Par définition, $\mathbf{y}_\ell$ s'écrit $\mathbf{y}_\ell = \mathbf{C}_\ell^{1/2}\mathbf{x}_\ell$ où la matrice $\mathbf{C}_\ell$ est définie dans l'énoncé du théorème 2, $\mathbf{x}_\ell$ est un vecteur à éléments iid centrés et $\mathbf{x}_\ell$ et $\mathbf{Q}_\ell$ sont indépendants. Par conséquent, $\frac{1}{t}\mathbf{y}_l^H\mathbf{Q}_\ell\mathbf{y}_\ell$ est proche de $\frac{1}{t}\mathrm{tr}(\mathbf{C}_\ell\mathbf{Q}_\ell)$ pour t grand. Par ailleurs, comme

$$(\mathbb{E}_\ell - \mathbb{E}_{\ell+1}) \log\left(1 + \frac{1}{t}\mathrm{tr}(\mathbf{C}_\ell\mathbf{Q}_\ell)\right) = 0 ,$$

l'équation (3) se réécrit $W_\ell = (\mathbb{E}_\ell - \mathbb{E}_{\ell+1})\log(1 + \gamma_\ell)$ où

$$\gamma_\ell = \frac{\frac{1}{t}\left(\mathbf{y}_\ell^H\mathbf{Q}_\ell\mathbf{y}_\ell - \mathrm{tr}(\mathbf{C}_\ell\mathbf{Q}_\ell)\right)}{1 + \frac{1}{t}\mathrm{tr}(\mathbf{C}_\ell\mathbf{Q}_\ell)}$$

est une petite quantité pour t grand. En utilisant l'approximation $\log(1+\gamma_\ell) \approx \gamma_\ell$ et en constatant que $\mathbb{E}_{\ell+1}\gamma_\ell = 0$, nous avons finalement

$$W_\ell \approx \mathbb{E}_\ell\gamma_\ell .$$

Afin d'établir le TLC, nous travaillons sur la somme d'incréments de martingales $\sum_{\ell=1}^t \mathbb{E}_\ell\gamma_\ell$. Les propriétés des résolvantes $\mathbf{Q}_\ell$ ainsi que leurs liens avec la matrice $\mathbf{G}$ définie dans l'énoncé du théorème 2 jouent un rôle fondamental dans cette étude.

L'utilisation de la méthode des martingales pour l'étude des fluctuations de fonctionnelles de matrices aléatoires a été initiée par Girko. Cette technique est également utilisée dans [2].

Références

[1] G.W. Anderson and O. Zeitouni, "A CLT for a band matrix model", *PTRF*, 134 (2), pp 283–338, 2006.

[2] Z.D. Bai and J.W. Silverstein, "CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices", *Ann. Prob.*, 32 (1A), pp 553–605, 2004.

[3] P. Billingsley, "Probability and Measure", Wiley, 1995.

[4] C.N. Chuah, D.N.C. Tse, J.M. Kahn, and R.A. Valenzuela, "Capacity Scaling in MIMO Wireless Systems Under Correlated Fading", *IEEE Trans. on IT*, 48 (3), pp 637–650, Mar. 2002.

[5] V.L. Girko, "Theory of Random Determinants", Volume 45 of Mathematics and its Applications (Soviet Series), Kluwer Academic Publishers Group, 1990.

[6] W. Hachem, O. Khorunzhiy, Ph. Loubaton, J. Najim and L. Pastur, "A New Approach for Capacity Analysis of Large Dimensional Multi-Antenna Channels", submitted to *IEEE Trans. on IT*, Nov. 2006, available online at `arXiv` portal.

[7] W. Hachem, Ph. Loubaton, J. Najim, "Deterministic equivalents for certain functionals of large random matrices", *Annals of Applied Probabilty*, 17 (3), 2007.

[8] W. Hachem, Ph. Loubaton, J. Najim, "A CLT for Information Theoretic Statistics of Gram Random Matrices with a Given Variance Profile", submitted to *PTRF*, June 2007. Available online at `HAL` and `arXiV` portals.

[9] V.A. Marchenko and L.A. Pastur, "Distribution of Eigenvalues for Some Sets of Random Matrices", *Math. USSR – Sbornik*, 1 (4), pp 457–483, 1967.

[10] A.L. Moustakas, S.H. Simon and A.M. Sengupta, "MIMO Capacity Through Correlated Channels in the Presence of Correlated Interferers and Noise : A (Not So) Large N Analysis" *IEEE Trans. on IT*, 49 (10), pp 2545–2561, Oct 2003.

[11] A.M. Sayeed, "Deconstructing multiantenna fading channels", *IEEE Trans. on SP*, pp. 2563-2579, Oct 2002.

[12] I.E. Telatar, "Capacity of Multi-Antenna Gaussian Channels", Technical Memorandum, Bell Laboratories, Lucent Technologies, October 1995, published in *ETT*, 10 (6), pp. 585–595, Nov/Dec 1999.

[13] A.M. Tulino, S. Verdu, "Random Matrix Theory and Wireless Communications", in Foundations and Trends in Communications and Information Theory, vol. 1, pp. 1-182, Now Publishers, June 2004.