

Information mutuelle de canaux radio et opérateurs de Jacobi ergodiques

Walid HACHEM¹, Adrien HARDY², Shlomo SHAMAI (SHITZ)³

¹CNRS / LIGM (UMR 8049) Université Paris-Est Marne-la-Vallée, France

²Laboratoire Paul Painlevé, Université de Lille, France

³Technion, Israel Institute of Technology, Israël

walid.hachem@u-pem.fr, adrien.hardy@univ-lille.fr, sshlomo@ee.technion.ac.il

Résumé – L’information mutuelle de Shannon de canaux radio ergodiques est étudiée sous l’angle de la théorie des opérateurs ergodiques dans l’espace de Hilbert $\ell^2(\mathbb{Z})$. De plus, en supposant que le canal soit un processus de Markov ergodique, l’approximation à fort rapport signal à bruit de l’information mutuelle est considérée. Le régime des grandes dimensions est également étudié.

Abstract – Shannon’s mutual information of ergodic radio channels is studied in the light of the ergodic operator theory in the Hilbert space $\ell^2(\mathbb{Z})$. Assuming moreover that the channel is an ergodic Markov process, the high Signal to Noise Ratio behavior of the mutual information is assessed. Finally, the large dimensional regime is considered.

1 Introduction

Afin d’introduire notre problème, nous considérons le modèle de communication sans fil sur un canal sélectif en temps et en fréquence

$$y_n = \sum_{\ell=0}^L c_{n,\ell} s_{n-\ell} + v_n \quad (1)$$

où L est le degré du canal, où les processus (s_n) , (y_n) et (v_n) qu’on suppose scalaires dans cet exemple représentent respectivement le signal émis, le signal reçu et le bruit additif et où le vecteur $C_n = [c_{n,0}, \dots, c_{n,L}]^T \in \mathbb{C}^{L+1}$ est le vecteur des coefficients temporels du canal à l’instant n . Dans un environnement mobile, la suite des vecteurs (C_n) est souvent considérée comme un processus aléatoire ergodique. En ne considérant que l’hypothèse générique d’ergodicité, notre objectif est de calculer l’information mutuelle moyenne de Shannon de ce canal en le supposant connu du récepteur. Si on empile $n - m + 1$ éléments du signal reçu où m et n sont deux entiers relatifs tels que $n \geq m$, le modèle de transmission « vectorisé » devient $[y_m, \dots, y_n]^T = B_{m,n} [s_{m-L}, \dots, s_n]^T + [v_m, \dots, v_n]^T$ où

$$B_{m,n} = \begin{bmatrix} c_{m,L} & \cdots & c_{m,0} & & \\ & \ddots & & \ddots & \\ & & & c_{n,L} & \cdots & c_{n,0} \end{bmatrix}.$$

Sous des hypothèses classiques sur les statistiques de (s_n) et de (v_n) (cf. plus loin) cette information mutuelle s’écrit alors

$$\mathcal{I}_\rho = \lim_{n-m \rightarrow \infty} \frac{\log \det(\rho B_{m,n} B_{m,n}^* + I_{n-m+1})}{n - m + 1} \quad (2)$$

où $\rho > 0$ est le rapport signal à bruit (RSB) et où limps désigne la limite presque sûre, limite dont l’existence découle de l’ergodicité de (C_n) . L’étude de \mathcal{I}_ρ s’insère dans le cadre théorique des opérateurs ergodiques sur l’espace de Hilbert $\ell^2(\mathbb{Z})$ couramment étudiés dans le domaine de la physique statistique [4]. Dans notre cas, un tel opérateur est l’analogie « doublement infini » de la matrice $B_{m,n}$ ci-dessus que l’on note avec un léger abus $B = B_{-\infty, \infty}$. La convergence (2) provient du fait que grâce à l’ergodicité, la mesure spectrale de la matrice $B_{m,n} B_{m,n}^*$ converge quand $n - m \rightarrow \infty$ vers une mesure de probabilité déterministe appelée densité d’états de l’opérateur ergodique BB^* . En physique statistique, l’étude de la densité d’états s’est focalisée le plus souvent sur les opérateurs de Jacobi (ou tridiagonaux) associés à l’équation de Schrödinger discrète. Pour de tels opérateurs, la densité d’état peut être caractérisée par la formule dite de Thouless [4].

Dans notre contexte, l’utilisation de la formule de Thouless s’avère toujours possible quand on considère l’opérateur BB^* ci-dessus comme un opérateur tridiagonal par blocs (ou bloc-Jacobi). Cette voie a été suivie par [3].

Dans ce travail, nous empruntons un autre chemin pour calculer \mathcal{I}_ρ . Toujours dans le cadre des opérateurs bloc-Jacobi, en exhumant un processus ergodique couplé avec le processus du canal, nous obtenons une expression différente et plus facilement manipulable que l’expression de \mathcal{I}_ρ obtenue par l’intermédiaire de la formule de Thouless. Nous exploitons l’expression obtenue selon deux directions : le régime à fort RSB pour un canal markovien et le régime des grandes dimensions selon lequel la taille des blocs converge vers l’infini.

Les preuves des résultats de cet article se trouvent dans [1].

2 Problème et résultat principal

Nous débutons par l'énoncé rigoureux de notre modèle de communication. La structure bloc-Jacobi de l'opérateur qui représente le canal sera mise en avant. Etant donné deux entiers positifs N et K , nous considérons le modèle de communication sans fil

$$Y_n = F_n S_{n-1} + G_n S_n + V_n \quad (3)$$

où les processus $(Y_n)_{n \in \mathbb{Z}}$, $(S_n)_{n \in \mathbb{Z}}$ et $(V_n)_{n \in \mathbb{Z}}$ représentent respectivement le signal reçu, le signal émis et le bruit additif et où $(F_n, G_n)_{n \in \mathbb{Z}}$ pour $F_n, G_n \in \mathbb{C}^{N \times K}$ est une représentation matricielle d'un canal de communication radio.

Le canal de communication scalaire décrit dans l'introduction est un cas particulier de ce modèle : si $L > 0$, nous posons $N = K = L$,

$$Y_n = \begin{bmatrix} y_{nL} \\ \vdots \\ y_{nL+L-1} \end{bmatrix}, S_n = \begin{bmatrix} s_{nL} \\ \vdots \\ s_{nL+L-1} \end{bmatrix}, V_n = \begin{bmatrix} v_{nL} \\ \vdots \\ v_{nL+L-1} \end{bmatrix},$$

$$F_n = \begin{bmatrix} c_{nL,L} & \cdots & c_{nL,1} \\ & \ddots & \vdots \\ & & c_{nL+L-1,L} \end{bmatrix} \in \mathbb{C}^{L \times L} \text{ et} \quad (4)$$

$$G_n = \begin{bmatrix} c_{nL,0} & & \\ \vdots & \ddots & \\ c_{nL+L-1,L-1} & \cdots & c_{nL+L-1,0} \end{bmatrix} \in \mathbb{C}^{L \times L}.$$

Si $L = 0$, nous posons $N = K = 1$, $Y_n = y_n$, $S_n = s_n$, $V_n = v_n$, $F_n = 0$ et $G_n = c_{n,0}$. On peut également considérer la variante MIMO de ce modèle où les coefficients $c_{n,\ell}$ deviennent des matrices $R \times T$ où R (resp. T) est le nombre d'antennes en réception (resp. en émission). Dans ce cas, si $L > 0$, les matrices F_n et G_n à valeurs dans $\mathbb{C}^{N \times K}$ deviennent des matrices bloc triangulaires comme en (4) avec $N = RL$ et $K = TL$.

Dans l'équation (1) ou dans sa variante MIMO, le paramètre n peut également représenter l'espace. Supposons que les stations de base (SB) d'un réseau cellulaire soient disposées sur un axe et que chaque SB reçoive les signaux de $L + 1$ utilisateurs qui lui sont proches. Comme ces utilisateurs communiquent également avec d'autres SB, le signal y_n reçu par la SB n à un moment donné est décrit par l'équation (1) où le temps est maintenant omis, où s_ℓ est le signal émis par l'utilisateur ℓ et où $c_{n,\ell}$ est le canal montant reliant l'utilisateur $n - \ell$ à la SB n .

Notre objectif est d'étudier l'information mutuelle de Shannon de ce canal. Pour ceci, nous nous plaçons dans le cas où (S_n) est un processus iid de loi $\mathcal{CN}(0, I_K)$, le bruit (V_n) est iid de loi $\mathcal{CN}(0, \rho^{-1} I_N)$ où $\rho > 0$ représente le RSB et les processus (S_n) , (V_n) et (F_n, G_n) sont indépendants. Notre hypothèse de base sera la suivante :

Hypothèse 1. *Le processus $(F_n, G_n)_{n \in \mathbb{Z}}$ est stationnaire et ergodique. De plus, $\mathbb{E}\|F_0\|^2 < \infty$ et $\mathbb{E}\|G_0\|^2 < \infty$.*

La stationnarité et l'ergodicité sont entendues dans le sens classique de la théorie ergodique, à savoir, dans l'espace de probabilité canonique du processus $(F_n, G_n)_{n \in \mathbb{Z}}$, l'opérateur de décalage $T\omega = (\dots, (F_0, G_0), (F_1, G_1), (F_2, G_2), \dots)$ pour $\omega = (\dots, (F_{-1}, G_{-1}), (F_0, G_0), (F_1, G_1), \dots)$ préserve la mesure de probabilité de ce processus et ne peut laisser invariants que des ensembles de probabilité 0 ou 1.

Si, pour deux entiers relatifs $m \leq n$, on se donne la matrice

$$H_{m,n} = \begin{bmatrix} F_m & G_m & & & \\ & F_{m+1} & G_{m+1} & & \\ & & \ddots & \ddots & \\ & & & F_n & G_n \end{bmatrix},$$

alors l'information mutuelle par composante du modèle de communication (3) est

$$\mathcal{I}_\rho = \limsup_{n-m \rightarrow \infty} \frac{\log \det (\rho H_{m,n} H_{m,n}^* + I)}{(n-m+1)N}$$

où la limite presque sûre existe grâce à l'hypothèse 1 [2] (remarquons ici que dans le cas cellulaire montant décrit ci-dessus, \mathcal{I}_ρ est l'information mutuelle - somme dans le cas où les SB collaborent entre elles mais non les mobiles). Cette limite est identifiée par le théorème suivant. La clé réside dans l'analyse de certaines résolvantes associées à l'opérateur bloc-Jacobi ergodique HH^* où H est l'analogue doublement infini $H_{-\infty, \infty}$ de $H_{m,n}$.

Dans la suite, \mathcal{H}_K^{++} (resp. \mathcal{H}_K^+) désigne le cône des matrices hermitiennes définies positives (resp. semi définies positives) dans $\mathbb{C}^{K \times K}$.

Théorème 1. *Si l'hypothèse 1 est satisfaite, alors :*

- (a) *Il existe un processus stationnaire unique $(W_n)_{n \in \mathbb{Z}}$ à valeurs dans \mathcal{H}_K^{++} et qui satisfait la récursion*

$$W_n = \left(I + \rho G_n^* (I + \rho F_n W_{n-1} F_n^*)^{-1} G_n \right)^{-1}.$$

En particulier, (W_n) est ergodique.

- (b) *L'information mutuelle par composante s'écrit*

$$\mathcal{I}_\rho = \frac{1}{N} \left(\mathbb{E} \log \det (I + \rho F_0 W_{-1} F_0^*) - \mathbb{E} \log \det W_0 \right).$$

- (c) *Soit le processus $(X_n)_{n \in \mathbb{N}}$ défini par la récursion*

$$X_n = \left(I + \rho G_n^* (I + \rho F_n X_{n-1} F_n^*)^{-1} G_n \right)^{-1}$$

où $X_{-1} \in \mathcal{H}_K^+$ est arbitraire. Alors

$$\mathcal{I}_\rho = \limsup_{n \rightarrow \infty} \frac{1}{nN} \sum_{\ell=0}^{n-1} \log \det (I + \rho F_\ell X_{\ell-1} F_\ell^*) - \log \det X_\ell. \quad (5)$$

3 Canal markovien et grand RSB

Dans le régime où $\rho \rightarrow \infty$, il arrive souvent que

$$\mathcal{I}_\rho = \min(K/N, 1) \log \rho + \kappa_\infty + o(1).$$

Notre objectif est d'établir cette convergence et de donner l'expression du décalage κ_∞ dans le cas où le processus (F_n, G_n) est un processus de Markov qui satisfait l'hypothèse de régularité suivante que nous énonçons dans le cas où $N > K$:

Hypothèse 2. Le processus stationnaire (F_n, G_n) est un processus de Markov qui admet une unique mesure de probabilité invariante. De plus,

(a) Ce processus de Markov est fellerien, à savoir, la fonction φ définie par l'équation

$$\varphi(F, G) = \mathbb{E}[f(F_{n+1}, G_{n+1}) \mid (F_n, G_n) = (F, G)]$$

est continue (et bornée) dès que la fonction f est continue et bornée,

(b) $\mathbb{E}\|F_0\|^2 + \mathbb{E}\|G_0\|^2 < \infty$,

(c) $\mathbb{E}|\log \det(F_0^* F_0)| < \infty$,

(d) Pour tout vecteur déterministe $v \in \mathbb{C}^K \setminus \{0\}$, nous avons

$$\det(G_n^* F_n) \neq 0 \text{ p.s.} \quad \text{et} \quad \Pi_{G_n}^\perp F_n v \neq 0 \text{ p.s.} \quad (6)$$

où Π_G^\perp est le projecteur orthogonal sur le sous-espace orthogonal aux colonnes de G .

Comme un processus de Markov ayant une unique mesure invariante est ergodique, l'hypothèse 2 implique l'hypothèse 1.

Théorème 2. Supposons $N > K$. Sous l'hypothèse 2,

(a) Il existe un unique processus stationnaire $(Z_n)_{n \in \mathbb{Z}}$ sur \mathcal{H}_K^{++} décrit par la récursion

$$Z_n = G_n^* (I + F_n Z_{n-1}^{-1} F_n^*)^{-1} G_n,$$

(b) La variable aléatoire $\log \det(Z_0 + F_1^* F_1)$ est intégrable et

$$\mathcal{I}_\rho = \frac{K}{N} \log \rho + \kappa_\infty + o_{\rho \rightarrow \infty}(1)$$

où

$$\kappa_\infty = \frac{1}{N} \mathbb{E} \log \det(Z_0 + F_1^* F_1).$$

(c) Pour tout $X_{-1} \in \mathcal{H}_K^{++}$, soit le processus $(X_n)_{n \in \mathbb{N}}$ défini récursivement par l'équation

$$X_n = G_n^* (I + F_n X_{n-1}^{-1} F_n^*)^{-1} G_n.$$

Alors

$$\kappa_\infty = \lim_{n \rightarrow \infty} \frac{1}{nN} \sum_{\ell=0}^{n-1} \log \det(X_\ell + F_{\ell+1}^* F_{\ell+1})$$

en probabilité.

Le cas où $N < K$ peut être traité d'une manière similaire. Le cas où $N = K$, où structurellement, l'hypothèse 2-(d) ne peut pas être satisfaite, ne sera pas considéré dans ce travail.

Il est utile de produire un modèle de canal MIMO pour lequel l'hypothèse 2 soit satisfaite. Un modèle AR pour le canal $C_n = [c_{n,0}^\top, \dots, c_{n,L}^\top]^\top \in \mathbb{C}^{R \times (L+1)T}$ sera considéré dans la proposition suivante.

Proposition 1. Etant donné trois entiers strictement positifs L, R et T tels que $R > T$, soit $(C_n)_{n \in \mathbb{Z}}$ le processus aléatoire sur $\mathbb{C}^{(L+1)R \times T}$ défini par la récursion

$$C_n = \begin{bmatrix} A_0 & & \\ & \ddots & \\ & & A_L \end{bmatrix} C_{n-1} + U_n \triangleq AC_{n-1} + U_n$$

où les $\{A_\ell\}_{\ell=0}^L$ sont des matrices déterministes $R \times R$ dont les spectres sont inclus dans le disque-unité ouvert, et où $(U_n)_{n \in \mathbb{Z}}$ est un processus matriciel iid tel que $\mathbb{E}\|U_0\|^2 < \infty$. Soit F_n et G_n les matrices $LR \times LT$ définies dans (4). Supposons que les éléments de U_n soient indépendants et que leurs mesures de probabilité soient absolument continues par rapport à la mesure de Lebesgue sur \mathbb{C} . Supposons enfin que leurs densités soient bornées. Alors, le processus (F_n, G_n) satisfait l'hypothèse 2.

Le modèle décrit dans cette proposition rend compte de la sélectivité en fréquence, de la sélectivité en temps, de la présence d'un profil de puissance des coefficients matriciels du canal et des corrélations au sein de ces coefficients.

4 Illustrations numériques

Dans ce paragraphe, nous illustrons les résultats des théorèmes 1 et 2. Nous considérons également le régime asymptotique où les dimensions N et K tendent vers l'infini au même rythme.

Pour commencer, nous considérons la variante MIMO du canal décrit par l'équation (1). Pour modéliser la sélectivité en temps du canal, nous recourons au modèle AR $c_{n,\ell} = \alpha c_{n-1,\ell} + \sqrt{1 - \alpha^2} a_\ell u_{n,\ell}$ où $\alpha \in [0, 1]$. En écrivant $\alpha = \exp(-f_d)$, le paramètre $f_d > 0$ est proportionnel à la fréquence Doppler. Pour $n \in \mathbb{Z}$, $\ell \in [L] \triangleq \{0, \dots, L\}$, $r \in [R]$ et $t \in [T]$, les coefficients $[u_{n,\ell}]_{r,t}$ sont iid de loi $\mathcal{CN}(0, T^{-1})$. Le vecteur réel $a = [a_0, \dots, a_L]$ représente le profil d'amplitude des coefficients de la réponse impulsionnelle.

Sur la figure 1, nous comparons l'estimée de l'information mutuelle $\widehat{\mathcal{I}}^{(m)} = (mN)^{-1} \log \det(\rho H_{1,m} H_{1,m}^* + I)$ à celle fournie par le théorème 1-(c). La première est représentée sur la figure 1 sous la forme de boîtes à moustaches obtenues pour 150 réalisations de la matrice $H_{1,m}$. Sur cette figure, la ligne horizontale à $\simeq 1.3$ nats représente la seconde en prenant $n = 4000$ dans le membre de droite de (5). Cette figure montre que

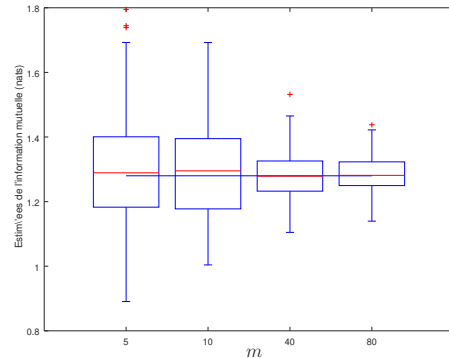


FIGURE 1 – $\widehat{\mathcal{I}}^{(m)}$ vs. m et estimée du Th. 1-(c). Paramètres : $R = T = 2$, $L = 3$, $f_d = 0.05$, $a_\ell \propto \exp(-0.4\ell)$, RSB = 6 dB.

les paramètres de dispersion de l'estimée $\widehat{\mathcal{I}}^{(m)}$ demeurent importants même pour $m = 80$, valeur pour laquelle $H_{1,m}H_{1,m}^* \in \mathbb{C}^{480 \times 480}$. Les estimées produites par le théorème 1 réclament beaucoup moins de calculs car elles ne mettent en jeu que l'inversion de matrices positives 6×6 .

La figure 2 illustre le comportement du paramètre κ_∞ prédit par le théorème 2. Dans cette expérience, nous considérons un modèle de canal plus général que le précédent en remplaçant les coefficients centrés $c_{n,\ell}$ de ce dernier par

$$\sqrt{\frac{K_R}{K_R + 1}} d_{n,\ell} + \sqrt{\frac{1}{K_R + 1}} c_{n,\ell},$$

où l'élément (r, t) de la matrice déterministe $d_{n,\ell}$ est de la forme $a_\ell \exp(2i\pi(r-t) \sin(\pi\ell/L))$ et où les poids relatifs de la partie déterministe et aléatoire sont commandés par le facteur de « Rice » $K_R \geq 0$. Remarquons que la proposition 1 s'applique à ce canal. La figure 2 illustre le comportement de κ_∞ en fonction de K_R et de la fréquence Doppler f_d . Cette simu-

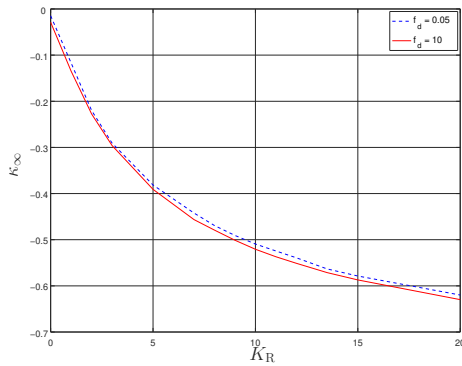


FIGURE 2 – comportement de κ_∞ . Paramètres : $R = 3, T = 2, L = 3, a_\ell \propto \exp(-0.4\ell)$.

lation réalisée à l'aide du théorème 2–(c) montre que l'impact de f_d est marginal pour notre modèle. Concernant K_R , la part d'aléa dans ce canal a un effet bénéfique en supposant, nous le rappelons, que le canal soit parfaitement connu au récepteur.

Pour finir, nous étudions le comportement de \mathcal{I}_ρ dans le cas où N et K tendent vers l'infini au même rythme. En revenant aux équations (4), ce régime peut correspondre au cas où $L \rightarrow \infty$ alors que R et T sont fixes, au cas où L est fixe alors que R et T tendent vers l'infini au même rythme ainsi qu'aux cas intermédiaires. Pour une grande classe de modèles statistiques du processus (F_n, G_n) , nous savons que dans ce régime asymptotique, la densité d'états de l'opérateur HH^* que l'on paramètre maintenant par N converge vers une mesure de probabilité prédite par le régime des grandes matrices aléatoires [2]. Une illustration de ce phénomène est fournie par la proposition suivante dont la preuve que nous détaillons dans [1] s'appuie sur le théorème 1.

Proposition 2. Fixons $T = R$ et supposons que L varie. Supposons que $c_{n,\ell} = a_\ell u_{n,\ell}$ où les scalaires a_ℓ sont tels que $\sup_L \max_\ell \sqrt{L}|a_\ell| < \infty$ et $\sum_{\ell=0}^L a_\ell^2 = 1$. Supposons que les

matrices $\{u_{n,\ell}, n \in \mathbb{Z}, \ell \in [L]\}$ soient iid et que les éléments de chacune de ces matrices soient iid, centrés et de variance $1/T$. Alors l'information mutuelle par composante satisfait

$$\mathcal{I}_\rho \xrightarrow{L \rightarrow \infty} 2 \log \frac{\sqrt{4\rho + 1} + 1}{2} - \frac{2\rho + 1 - \sqrt{4\rho + 1}}{2\rho}. \quad (7)$$

Le terme de droite de (7) coïncide en réalité avec $\int \log(1 + \rho\lambda) \mu_{\text{MP}}(d\lambda)$ où μ_{MP} est une loi de Marchenko-Pastur, bien connue dans le domaine des matrices aléatoires.

Afin d'illustrer ce résultat, nous revenons au modèle cellulaire introduit ci-dessus et nous supposons que le réseau de SB devient de plus en plus dense. De ce fait, le nombre L d'interfereurs qui communiquent avec une SB augmente linéairement avec le nombre de SB. Dans le cas où les SB collaborent entre elles, la proposition 2 montre alors que l'information mutuelle par unité de distance croît linéairement avec L . Dans la figure 3, une décroissance cubique des puissances reçues en fonction de la distance est considérée. L'approximation du ré-

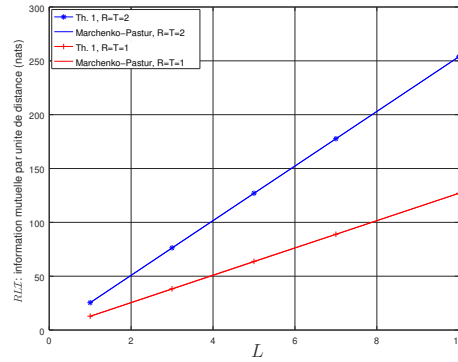


FIGURE 3 – information mutuelle par unité de distance vs. densité des SB. $\rho = 6$ dB.

gime des grandes matrices aléatoires est précise pour L aussi petit que 1.

Remerciements. Le travail des deux derniers auteurs a été partiellement soutenu par les projets BoB et par le programme européen Horizon 2020, projet no. 694630.

Références

- [1] W. Hachem, A. Hardy, and S. Shamai. Mutual information of wireless channels and block-Jacobi ergodic operators. *arXiv:1811.04734*, 2018.
- [2] W. Hachem, A. M. Moustakas, and L. Pastur. The Shannon's mutual information of a multiple antenna time and frequency dependent channel : An ergodic operator approach. *J. Math. Phys.*, 56(11), 2015.
- [3] N. Levy, O. Zeitouni, and S. Shamai. On information rates of the fading Wyner cellular model via the Thouless formula for the strip. *IEEE Trans. IT*, 56(11), 2010.
- [4] L. Pastur and A. Figotin. *Spectra of random and almost-periodic operators*. Springer-Verlag, 1992.