

Sur l'estimation de la capacité ergodique des grands systèmes MIMO

Pascal VALLET, Philippe LOUBATON

IGM LabInfo, UMR-CNRS 8049
Université Paris-Est
5, Boulevard Descartes, Champs-sur-Marne
77454 Marne-la-Vallée Cedex 2, France
{vallet, loubaton}@univ-mlv.fr

Résumé – Dans cet article, nous nous intéressons à l'estimation de la capacité ergodique des grands systèmes MIMO à partir de séquences d'apprentissage dont la longueur est du même ordre de grandeur que le nombre d'antennes du système. Dans ce contexte, les techniques standards d'estimation ont de mauvaises performances. En utilisant la théorie de la G-estimation initiée par Girko, nous proposons un nouvel estimateur et analysons son comportement asymptotique. Des résultats numériques montrent qu'il surpasse de façon significative les performances de l'estimateur standard pour un nombre raisonnable d'antennes.

Abstract – In this paper, we deal with the estimation of the ergodic capacity of large MIMO systems, using training sequences whose lengths are of the same order of magnitude than the number of antennas. In this context, the traditional estimator becomes inconsistent. Following the ideas developed by Girko in the context of the so-called theory of G-estimation, we propose a new estimator. We analyze its asymptotic behaviour and show, using numerical simulations, that it tends to improve significantly the performance of the standard estimate for a realistic number of antennas.

1 Introduction

Pour prédire les performances potentielles des systèmes de communications numériques, il est utile d'avoir accès aux propriétés statistiques des canaux de propagation car elles permettent d'évaluer divers indicateurs de performances. Pour cela, des techniques dites de "sondage de canal" sont utilisées. Dans cet article, nous nous intéressons à l'estimation de la capacité ergodique des systèmes MIMO mobiles dans ce contexte de sondage de canal. Pour estimer le canal, on utilise couramment une séquence d'apprentissage de longueur suffisante. Cependant, dans un contexte de mobilité, les séquences d'apprentissage utilisées doivent être de taille réduite, car le canal doit rester constant durant leur transmission. Si la taille L de la séquence d'apprentissage est du même ordre de grandeur que le nombre d'antennes du système, l'estimateur standard de la capacité a de piètres performances en raison du bruit additif dont la contribution n'est pas négligeable. Dans ce contexte, nous pouvons mentionner les récents travaux de [4] dans lesquels un estimateur de la capacité est établi à l'aide de la théorie des probabilités libres. Cependant, cet estimateur n'est pas consistant dans le cas de grands systèmes MIMO.

Pour mettre au point un nouvel estimateur dans ce contexte, nous utilisons dans cet article la théorie des grandes matrices aléatoires, et faisons l'hypothèse que le nombre d'antennes en émission et réception, N et M , ainsi que la longueur de la séquence d'apprentissage L , convergent vers $+\infty$ au même rythme. Dans ce régime asymptotique, nous proposons un nouvel estimateur de la capacité ergodique basé sur la théorie de la G-

estimation, initiée par Girko.

Cet article est organisé comme suit. En section 2, nous présentons le modèle des signaux, le contexte précis de l'estimation et l'énoncé du problème. En section 3, nous donnons quelques résultats utiles en théorie des grandes matrices aléatoires. En section 4, nous proposons un nouvel estimateur consistant de la capacité ergodique du canal MIMO. En section 5, nous donnons quelques résultats numériques qui illustrent les précédents résultats.

2 Modèle et énoncé du problème

Nous considérons un système de mesure du canal MIMO constitué de N antennes d'émission et de M antennes de réception, et nous supposons que le canal est à évanouissements par blocs. Pour estimer la capacité ergodique du canal, l'émetteur envoie périodiquement une séquence d'apprentissage de taille L , notée $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_L]$, avec $\mathbf{s}_1, \dots, \mathbf{s}_L$ des vecteurs de \mathbb{C}^N tels que $\mathbf{S}\mathbf{S}^\dagger = L\mathbf{I}_N$. La matrice du canal est supposée constante sur chaque "slot", i.e durant la transmission de chaque séquence d'apprentissage. Nous notons T le nombre de slots disponibles, et nous supposons les propriétés statistiques du canal invariantes durant la transmission de ces T slots. Si $\mathbf{H}_{(t)}$ ($t = 1, \dots, T$) représente la matrice du canal durant le slot t , normalisée de telle sorte que $\mathbb{E}[\frac{1}{M}\text{Tr} \mathbf{H}_{(t)}\mathbf{H}_{(t)}^\dagger] = 1$, le

problème consiste alors à estimer

$$C(\sigma^2) = \frac{1}{N} \frac{1}{T} \sum_{t=1}^T \log \det \left(\mathbf{I}_M + \frac{\mathbf{H}_{(t)} \mathbf{H}_{(t)}^\dagger}{\sigma^2} \right) \quad (1)$$

Cette dernière quantité est clairement une approximation de la capacité ergodique (par antenne d'émission). Pour $t = 1, \dots, T$, la matrice des observations $\mathbf{Y}_{(t)} \in \mathcal{M}_{M,L}(\mathbb{C})$ correspondant à la transmission de la séquence \mathbf{S} est donnée par

$$\mathbf{Y}_{(t)} = \mathbf{H}_{(t)} \mathbf{S} + \mathbf{V}_{(t)} \quad (2)$$

avec $\mathbf{V}_{(t)}$ la matrice due au bruit, dont les éléments sont iid, gaussiens, centrés et de variance σ_v^2 . L'estimée $\mathbf{Z}_{(t)}$ du canal $\mathbf{H}_{(t)}$, basée sur la séquence d'apprentissage, est définie par

$$\mathbf{Z}_{(t)} = \frac{1}{L} \mathbf{Y}_{(t)} \mathbf{S}^\dagger = \mathbf{H}_{(t)} + \frac{\rho}{\sqrt{N}} \mathbf{W}_{(t)} \quad (3)$$

avec $\rho = \sqrt{\frac{\sigma_v^2 N}{L}}$ et $\mathbf{W} \in \mathcal{M}_{M,N}(\mathbb{C})$ une matrice aléatoire telle que $\mathbf{W}_{ij} \sim \mathcal{CN}(0, 1)$. Pour garder ρ constant, nous supposons dans cet article que la taille de la séquence d'apprentissage L augmente linéairement avec N .

Une manière classique d'estimer (1) consiste à remplacer $\mathbf{H}_{(t)} \mathbf{H}_{(t)}^\dagger$ par son estimée standard non biaisée $\mathbf{Z}_{(t)} \mathbf{Z}_{(t)}^\dagger - \rho^2 \mathbf{I}_M$. Cependant, certaines valeurs propres de $\mathbf{Z}_{(t)} \mathbf{Z}_{(t)}^\dagger$ peuvent être inférieures à ρ^2 . En notant $\hat{\lambda}_1^{(t)}, \dots, \hat{\lambda}_M^{(t)}$ les valeurs propres de $\mathbf{Z}_{(t)} \mathbf{Z}_{(t)}^\dagger$ (dans l'ordre croissant), l'estimateur standard de (1) est donné par

$$\hat{C}_{\text{trad}}(\sigma^2) = \frac{1}{N} \frac{1}{T} \sum_{t=1}^T \sum_{m=1}^M \log \left(1 + \frac{(\hat{\lambda}_m^{(t)} - \rho^2)^+}{\sigma^2} \right) \quad (4)$$

avec $(\cdot)^+$ la partie positive. Cet estimateur est pertinent dans le cas où $L \rightarrow +\infty$ et $N/L \rightarrow 0$, mais ce régime n'est pas réaliste dans le contexte des systèmes mobiles. Nous supposons donc que $N \rightarrow +\infty$, $L \rightarrow +\infty$ de telle sorte que $\rho = \sqrt{\frac{\sigma_v^2 N}{L}}$ reste constant. Nous supposons également que $M < N$, $M \rightarrow +\infty$, $N \rightarrow +\infty$ et que $c = \frac{M}{N}$ converge vers une constante $\alpha \in]0, 1[$. Dans ce contexte, (4) n'est plus consistant car $\|\mathbf{Z}_{(t)} \mathbf{Z}_{(t)}^\dagger - (\rho^2 \mathbf{I}_M + \mathbf{H}_{(t)} \mathbf{H}_{(t)}^\dagger)\|$ ne converge plus vers 0.

En utilisant des outils provenant de la théorie des grandes matrices aléatoires, nous pouvons mettre en évidence un estimateur consistant de (1) dans ce dernier contexte. Les résultats de cet article peuvent être étendus au cas où $c > 1$ en considérant la matrice $\mathbf{Z}^\dagger \mathbf{Z}$ en lieu et place de $\mathbf{Z} \mathbf{Z}^\dagger$.

3 Résultats préliminaires

Le modèle (3) est communément appelé "Modèle Information plus Bruit" et nous donnons dans cette section quelques résultats concernant les propriétés statistiques du spectre de la matrice aléatoire $\mathbf{Z} \mathbf{Z}^\dagger$ (la notation t qui traduit la dépendance au slot t est temporairement masquée pour cette section). On

note respectivement par $\lambda_1, \dots, \lambda_M$ et $\hat{\lambda}_1, \dots, \hat{\lambda}_M$ les valeurs propres (dans l'ordre croissant) des matrices $\mathbf{H} \mathbf{H}^\dagger$ et $\mathbf{Z} \mathbf{Z}^\dagger$.

Par souci de lisibilité, nous ne mentionnons pas explicitement la dépendance en M, N, L des quantités dépendantes de la matrice \mathbf{H} . Dans le reste de l'article, la mention " $N \rightarrow +\infty$ " sera équivalente à " $M, N, L \rightarrow +\infty$ tandis que $c \rightarrow \alpha \in]0, 1[$ et ρ reste constant". La notation " $\bar{X} \approx X$ " sera quant à elle équivalente à " $\bar{X} - X \rightarrow 0$ presque sûrement quand $N \rightarrow +\infty$ ", avec \bar{X} déterministe et X aléatoire.

Soit la mesure spectrale empirique $\hat{\mu}$ de la matrice aléatoire $\mathbf{Z} \mathbf{Z}^\dagger$ donnée par

$$\hat{\mu} = \frac{1}{M} \sum_{m=1}^M \delta_{\hat{\lambda}_m} \quad (5)$$

avec $\delta_{\hat{\lambda}_m}$ la mesure de Dirac au point $\hat{\lambda}_m$. On peut caractériser $\hat{\mu}$ par sa transformée de Stieltjes $\Upsilon_{\hat{\mu}}$ définie par

$$\Upsilon_{\hat{\mu}}(z) = \int_{\mathbb{R}^+} \frac{\hat{\mu}(d\lambda)}{\lambda - z} = \frac{1}{M} \text{Tr} (\mathbf{Z} \mathbf{Z}^\dagger - z \mathbf{I}_M)^{-1} \quad (6)$$

$\hat{\mu}$ est clairement aléatoire et le théorème fondamental qui suit (voir [1]) permet de caractériser son comportement asymptotique.

Théorème 1. *On suppose que $\sup_{M,N} \|\mathbf{H}\| < \infty$.*

Alors quand $N \rightarrow +\infty$, presque sûrement, $\hat{\mu}$ converge en distribution vers une mesure déterministe μ appelée mesure spectrale limite de $\mathbf{Z} \mathbf{Z}^\dagger$.

μ est caractérisée par la fonction $\delta(x) = \rho c \Upsilon_{\mu}(x)$ avec Υ_{μ} la transformée de Stieltjes de μ . $\forall x \in \mathbb{R}_^-$, $\delta(x)$ est la solution de l'équation*

$$\delta(x) = \frac{\rho}{N} \text{Tr} \mathbf{T}(x)$$

$$\mathbf{T}(x) =$$

$$\left[(-x(1 + \rho\delta(x)) + \rho^2(1 - c)) \mathbf{I}_M + \frac{\mathbf{H} \mathbf{H}^\dagger}{1 + \rho\delta(x)} \right]^{-1}$$

De plus, si $\mathcal{S} = \text{supp}(\mu)$ désigne le support de la mesure spectrale limite de $\mathbf{Z} \mathbf{Z}^\dagger$, alors $\delta(x)$ est holomorphe sur $\mathbb{R} - \mathcal{S}$.

On déduit de (6) et du théorème 1 que $\forall x \in \mathbb{R}_*^-$, $\Upsilon_{\hat{\mu}}(x) \approx \Upsilon_{\hat{\mu}}(x)$ et donc que

$$\delta(x) \approx \hat{\delta}(x) \quad (7)$$

avec $\hat{\delta}(x) = \frac{\rho}{N} \text{Tr} (\mathbf{Z} \mathbf{Z}^\dagger - x \mathbf{I}_M)^{-1}$. $\hat{\delta}(x)$ est donc un estimateur consistant de $\delta(x)$ si $x \in \mathbb{R}_*^-$.

Le théorème suivant (voir [5]) caractérise la borne inférieure de \mathcal{S} , le support de μ .

Théorème 2. *On se place dans le cas où $\mathbf{H} \mathbf{H}^\dagger$ est de rang plein. Soit $\phi(w)$ la fonction définie sur $\mathbb{R} - \{\lambda_1, \dots, \lambda_M\}$ par*

$$\phi(w) = w(1 - \rho^2 f(w))^2 + \rho^2(1 - c)(1 - \rho^2 f(w))$$

avec $f(w) = \frac{1}{N} \text{Tr} [\mathbf{H} \mathbf{H}^\dagger - w \mathbf{I}_M]^{-1}$. Soit w^- le plus petit antécédent des extrema locaux de ϕ , tel que $1 - \rho^2 f(w^-) > 0$ et $\phi(w^-) > 0$. Alors on a

$$x^- = \phi(w^-) = \inf \mathcal{S}$$

De plus, ϕ est une bijection strictement croissante de $]-\infty, w^-[$ dans $]-\infty, x^-[$. Soit $\psi(x)$ sa fonction réciproque sur $]-\infty, x^-[$. Alors on a

$$\psi(x) = x(1 + \rho\delta(x))^2 - \rho^2(1-c)(1 + \rho\delta(x))$$

et pour $w < w^-$, l'équation $\psi(x) = w$ admet une unique solution sur $]-\infty, x^-[$ égale à $\phi(w)$.

Les figures 1 et 2 représentent les deux allures possibles de ϕ au voisinage de 0 quand $c < 1$. Nous avons également représenté le point w_0 , solution de $\phi(w) = 0$ dans l'intervalle $]-\infty, w^-[$, qui sera utile pour la validité des calculs de performances de la section suivante. Notons que dans les deux cas, $0 \notin \mathcal{S}$, car $c < 1$. Nous donnons à présent un résultat fonda-

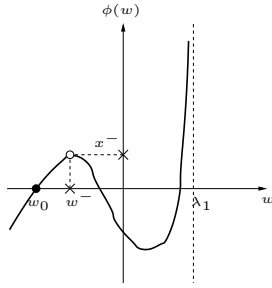


FIG. 1 – Allure de ϕ sur $]-\infty, \lambda_1[$ quand $c < 1$ et $f(0) > \frac{1}{\rho^2}$

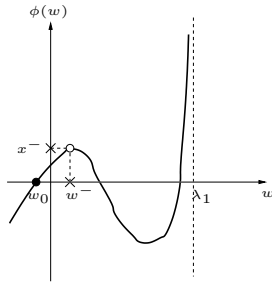


FIG. 2 – Allure de ϕ sur $]-\infty, \lambda_1[$ quand $c < 1$ et $f(0) < \frac{1}{\rho^2}$

mental sur les valeurs propres de $\mathbf{Z}\mathbf{Z}^\dagger$ (voir [6] (Theorem 23.1 p.267)).

Théorème 3. Pour tout $\epsilon > 0$, pour M, N suffisamment grands, la plus petite valeur propre $\hat{\lambda}_1$ de $\mathbf{Z}\mathbf{Z}^\dagger$ vérifie $\hat{\lambda}_1 > x^- - \epsilon$ presque sûrement.

Le théorème suivant se déduit du théorème 1 (voir [2] [3]).

Théorème 4. Soit la fonction \bar{C} définie sur \mathbb{R}_*^- par

$$\begin{aligned} \bar{C}(x) &= \frac{1}{N} \log \det \left(\mathbf{I}_M - \frac{\mathbf{H}\mathbf{H}^\dagger}{\psi(x)} \right) \\ &+ \log(1 + \rho\delta(x)) \\ &+ c \log \left(1 + \rho\delta(x) - \frac{\rho^2(1-c)}{x} \right) \\ &+ x\delta(x) \left(\delta(x) - \frac{\rho(1-c)}{x} \right) \end{aligned}$$

Alors,

$$\bar{C}(x) \approx \frac{1}{N} \log \det \left(\mathbf{I}_M - \frac{1}{x} \mathbf{Z}\mathbf{Z}^\dagger \right) \quad (8)$$

De plus, on a

$$\mathbb{E} \left[\frac{1}{N} \log \det \left(\mathbf{I}_M - \frac{1}{x} \mathbf{Z}\mathbf{Z}^\dagger \right) \right] = \bar{C}(x) + \mathcal{O} \left(\frac{1}{N^2} \right) \quad (9)$$

$$\text{Var} \left[\frac{1}{N} \log \det \left(\mathbf{I}_M - \frac{1}{x} \mathbf{Z}\mathbf{Z}^\dagger \right) \right] = \mathcal{O} \left(\frac{1}{N^2} \right) \quad (10)$$

4 Calcul du G-estimateur

Dans cette section, nous présentons un nouvel estimateur consistant de (1), basé sur le théorème 4. Pour la suite, nous reprenons la notation (t) pour toutes les quantités des précédentes sections qui dépendent de la matrice $\mathbf{H}_{(t)}$ (sur un slot t).

La construction repose sur l'utilisation conjointe d'une extension sur $]-\infty, x_{(t)}^-[$ des approximations (7) et (8) ainsi que sur le changement de variable $\psi(x) = -\sigma^2$ dans (8) qui permet de faire apparaître le terme $\frac{1}{N} \log \det \left(\mathbf{I}_M + \frac{\mathbf{H}\mathbf{H}^\dagger}{\sigma^2} \right)$ et de l'exprimer en fonction de quantités que l'on peut estimer de façon consistante.

Proposition 1. Quand $N \rightarrow +\infty$, $\forall \sigma^2 > -\min_{t=1, \dots, T} \{w_{(t)}^-\}$, $C(\sigma^2)$ donné par (1) est estimé de façon consistante par

$$\begin{aligned} \hat{C}_{\text{new}}(\sigma^2) &= \frac{1}{T} \sum_{t=1}^T \left[\frac{1}{N} \log \det \left(\mathbf{Z}_{(t)} \mathbf{Z}_{(t)}^\dagger - \hat{x}_*^{(t)} \mathbf{I}_M \right) \right. \\ &+ (c-1) \log \left(1 + \rho \hat{\delta}_{(t)}(\hat{x}_*^{(t)}) \right) \\ &\left. - \hat{x}_*^{(t)} \left(\hat{\delta}_{(t)}(\hat{x}_*^{(t)}) - \frac{\rho(1-c)}{\hat{x}_*^{(t)}} \right) \right] \\ &- c \log(\sigma^2) \end{aligned}$$

avec $\hat{x}_*^{(t)}$ la solution de l'équation $\hat{\psi}_{(t)}(x) = -\sigma^2$ et

$$\hat{\psi}_{(t)}(x) = x \left(1 + \rho \hat{\delta}_{(t)}(x) \right)^2 - \rho^2(1-c) \left(1 + \rho \hat{\delta}_{(t)}(x) \right)$$

La condition $\sigma^2 > -\min_{t=1, \dots, T} \{w_{(t)}^-\}$ apparaît car il est nécessaire que la solution de l'équation $\psi(x) = -\sigma^2$ soit inférieure à x^- pour garantir la consistance de l'estimateur. Nous caractérisons à présent les performances asymptotiques de l'estimateur en terme de biais et d'erreur quadratique moyenne (EQM). On rappelle que $w_{0,(t)}$ est la solution de l'équation $\phi_{(t)}(w) = 0$ dans l'intervalle $]-\infty, w_{(t)}^-[$.

Propriété 1. $\forall \sigma^2 > -\min_{t=1, \dots, T} \{w_{0,(t)}\}$ la quantité $\hat{C}_{\text{new}}(\sigma^2)$ définie en (11) est telle que

$$\mathbb{E} \left[\hat{C}_{\text{new}}(\sigma^2) \right] = C(\sigma^2) + \mathcal{O} \left(\frac{1}{N^2} \right) \quad (11)$$

$$\mathbb{E} \left[\left(\hat{C}_{\text{new}}(\sigma^2) - C(\sigma^2) \right)^2 \right] = \mathcal{O} \left(\frac{1}{N^2} \right) \quad (12)$$

Il faut remarquer que les précédents résultats du biais et de l'EQM ne sont a priori valides que si $\sigma^2 > -\min_{t=1,\dots,T}\{w_{0,(t)}\}$. Ils restent probablement vrais pour $\sigma^2 > -\min_{t=1,\dots,T}\{w_{(t)}^-\}$ car (11) et (12) sont sans doute vérifiés si $0 < x < x^-$, mais cela n'a pas encore été prouvé.

5 Résultats numériques

Dans cette section, nous comparons les performances de l'estimateur standard et du G-estimateur en terme de biais et d'EQM.

Les matrices de canal $\mathbf{H}_{(t)}$ suivent le "modèle de Kronecker", i.e, $\mathbf{H}_{(t)} = \frac{1}{\sqrt{N}} \mathbf{C}_{(t)}^{\frac{1}{2}} \mathbf{X} \tilde{\mathbf{C}}_{(t)}^{\frac{1}{2}}$, avec $\mathbf{C}_{(t)}$, $\tilde{\mathbf{C}}_{(t)}$ matrices réelles définies positive (et respectivement de tailles $M \times M$ et $N \times N$) et \mathbf{X} une matrice aléatoire telle que $\mathbf{X}_{ij} \sim \mathcal{CN}(0, 1)$. Pour les simulations, $\mathbf{C}_{(t)}$, $\tilde{\mathbf{C}}_{(t)}$ sont définies par

$$\mathbf{C}_{(t)} = [\eta_{1,(t)}^{|i-j|}]_{1 \leq i,j \leq M} \text{ et } \tilde{\mathbf{C}}_{(t)} = [\eta_{2,(t)}^{|i-j|}]_{1 \leq i,j \leq N}$$

$\eta_{1,(t)}$, $\eta_{2,(t)}$ étant des constantes comprises entre $] -1, 1[$.

Pour garder ρ constant dans (3), nous posons $L = 2N$. La résolution numérique de l'équation $\hat{\psi}_{(t)}(x) = -\sigma^2$ est obtenue en utilisant la méthode de Newton. De plus, les simulations ont été réalisées pour $T = 30$, $\sigma_v^2 = 0.25$ et $\sigma^2 = 1$.

Sur la figure 3, nous comparons les biais de (4) et (11) quand N, M augmentent ($c = 0.5$). On constate clairement que le biais de (4) reste constant avec N tandis que le biais de (11) décroît. Malheureusement, le rythme de convergence du biais en figure 3 n'est pas en $\frac{1}{N^2}$ pour des grandes valeurs de N . Cela est dû au fait que les valeurs du biais correspondantes sont du même ordre de grandeur que l'erreur commise lors de la résolution numérique de l'équation $\hat{\psi}(x) = -\sigma^2$.

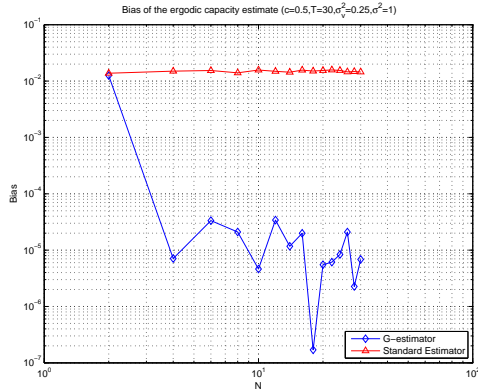


FIG. 3 – Comparaison des biais en fonction de N

La figure 4 représente l'évolution de l'EQM en fonction de N ($c = 0.5$). On constate que l'EQM de (4) est constante tandis que celle de (11) décroît clairement en $\frac{1}{N^2}$, ce qui confirme la propriété 1.

Les hypothèses sur le G-estimateur nous restreignent quant au choix de la valeur de σ^2 pour l'évaluation de la capacité. En effet, nous avons la condition $\sigma^2 > -\min_{t=1,\dots,T}\{w_{(t)}^-\}$

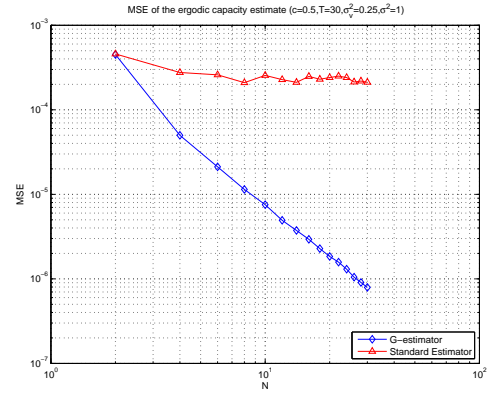


FIG. 4 – Comparaison des EQM en fonction de N

(où à $\sigma^2 > -\min_{t=1,\dots,T}\{w_{0,(t)}^-\}$ pour la validité du biais et de l'EQM). La figure 5 présente le tracé du minimum de w^- et w_0 en fonction de σ_v^2 , pour un grand nombre de tirages de matrices \mathbf{H} ($c = 0.5$ et $L = 2N$), et on voit clairement que la condition sur σ^2 n'est pas si restrictive dans le contexte des expérimentations présentes.

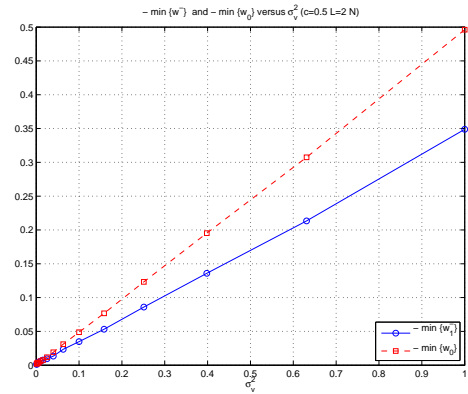


FIG. 5 – $-\min\{w^-\}$ et $-\min\{w_0\}$ en fonction de σ_v^2

Références

- [1] R.B. Dozier and J.W. Silverstein. On the empirical distribution of eigenvalues of large dimension information-plus-noise type matrices. *Journal of Multivariate Analysis* 98, pages 678–694, 2007.
- [2] J. Dumont, W. Hachem, S. Lasaulce, Ph. Loubaton, and J. Najim. On the capacity achieving covariance matrix for rician mimo channels : An asymptotic approach. *submitted to IEEE Transactions on Information Theory*. Available at <http://arxiv.org/abs/0710.4051>.
- [3] G.Taricco. Asymptotic mutual information statistics of separately correlated rician fading mimo channels. *IEEE Transactions on Information Theory*, 54(8):3490 – 3504, August 2008.
- [4] O.Ryan and M.Debbah. Channel capacity estimation using free probability theory. *IEEE Transactions on Signal Processing*, 56(11):5654–5667, November 2008.
- [5] P.Vallet, P.Loubaton, and X.Mestre. Improved subspace doa estimation with large arrays : The deterministic signal case. In *ICASSP, Taipei (Taiwan)*, 19-24 Apr. 2009. 4p.
- [6] V.L.Girko. *An Introduction to Statistical Analysis of Random Arrays*. VSP Science, December 1998.