

Grandes matrices aléatoires et systèmes MIMO.

Travaux en collaboration avec J. Dumont (UMLV), W. Hachem (Supelec), J. Najim (CNRS/ENST).

Plan de l'exposé.

- Introduction du type de problème applicatif : évaluation des performances de systèmes MIMO
- Introduction à l'analyse de l'information mutuelle en grandes dimensions
- le cas du modèle de Kronecker centré
- Le cas non centré

Articles à télécharger sur le site de l'ACI NIM MALCOM.

Les systèmes MIMO.

n antennes d'émissions, N antennes de réceptions.

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{v}$$

- $\mathbf{s}$ vecteur de dimension n dont les composantes représentent n symboles transmis (à un instant donné)
- $\mathbf{r}$ vecteur de dimension N reçu
- $\mathbf{v}$ bruit additif, $\mathbb{E}(\mathbf{v}\mathbf{v}^T) = \sigma^2\mathbf{I}$
- $\mathbf{H}$ matrice $N \times n$ des gains entre émetteurs et récepteurs

Le récepteur doit estimer $\mathbf{s}$ à partir de $\mathbf{r}$, $\mathbf{H}$ étant connue.

La problématique.

H réalisation d'un certain type de matrice aléatoire

Evaluer l'influence des propriétés statistiques de **H** sur les performances du système.

Exemples d'indices de performances usuels.

- Taux d'erreur en sortie d'un récepteur linéaire: si $\mathbf{H} = (\mathbf{h}_1, \mathbf{H}_2)$,
 $\mathbb{E}_{\mathbf{H}}(Q(\sqrt{\beta_1}))$ avec $\beta_1 = \mathbf{h}_1^T (\mathbf{H}_2 \mathbf{H}_2^T + \sigma^2 \mathbf{I})^{-1} \mathbf{h}_1$ et $Q(x)$ fonction d'erreur gaussienne.
- Information mutuelle moyenne: $\mathbb{E}_{\mathbf{H}} \left[\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H} \mathbf{H}^T}{\sigma^2} \right) \right]$
Débit maximum atteignable lorsque les n symboles transmis sont indépendants et de mêmes puissances.
- Capacité ergodique: $\text{Max}_{\mathbf{P} \geq 0, \frac{1}{n} \text{tr}(\mathbf{P}) \leq 1} \mathbb{E}_{\mathbf{H}} \left[\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H} \mathbf{P} \mathbf{H}^T}{\sigma^2} \right) \right]$
Débit maximum lorsque l'on optimise l'émetteur : $\mathbf{s} = \mathbf{P}_*^{1/2} \mathbf{s}_*$

Types d'hypothèses sur $\mathbf{H}$.

$\mathbf{H}$ matrice aléatoire gaussienne dans le contexte MIMO (mais non gaussienne dans d'autres contextes)

Différents scénarios considérés dans la littérature.

- $\mathbf{H} = \frac{1}{\sqrt{n}}\mathbf{X}$, $\mathbf{X}$ $N \times n$ iid centrée, variances 1.
- $\mathbf{H} = \frac{1}{\sqrt{n}}\mathbf{R}^{1/2}\mathbf{X}\tilde{\mathbf{R}}^{1/2}$, $\mathbf{X}$ $N \times n$ iid centrée, variances 1, $\mathbf{R} \geq 0$, $\tilde{\mathbf{R}} \geq 0$, modèle de Kronecker
- $\mathbf{H} = \frac{1}{\sqrt{n}}\mathbf{Z}$, $Z_{k,l}$ champs stationnaire centré : $\mathbb{E}(Z_{k,l}Z_{p,q}) = \phi(k-p, l-q)$
- Tous les modèles précédents dans le cas non centré : canaux de Rice.

Dans le cas gaussien, il est parfois possible de calculer les indices de performance, mais expressions inutilisables.

Etude des indices de performance en grandes dimensions.

- Etudier le comportement des indices de performances quand $N \rightarrow +\infty$,
 $n \rightarrow +\infty$, $\frac{N}{n} \rightarrow \alpha$
- Espérer que les expressions obtenues sont plus faciles à utiliser
- Espérer que les expressions obtenues sont de bonnes approximations pour des valeurs réalistes de N et de n .

Premiers articles publiés en 1999 dans la communauté de la théorie de l'information et des communications numériques : Tsé-Hanly et Verdu-Shamai.

Actuellement, une dizaine d'équipes travaillent sur cette approche.

ACI NIM Malcom labélisée en 2004, Eurecom (M. Debbah), Inst. de Physique des Basses Températures (Kharkov, L. Pastur), LTCI (ENST, J. Najim, Resp.), SUPELEC (W. Hachem), UMLV (P. Loubaton), UVSQ (O. Khorunzhy).

L'étude de l'information mutuelle moyenne I .

$$I_N(\sigma^2) = \mathbb{E}_{\mathbf{H}} \left[\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) \right]$$

Le schéma général.

- $\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right)$ est une fonctionnelle des valeurs propres $(\lambda_i)_{i=1,\dots,N}$ de la matrice de Gram $\mathbf{H}\mathbf{H}^T$ de $\mathbf{H}$:

$$\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) = \frac{1}{N} \sum_{i=1}^N \log \left(1 + \frac{\lambda_i}{\sigma^2} \right)$$

- Sous certaines hypothèses, la distribution empirique des valeurs propres de $\mathbf{H}\mathbf{H}^T$, $\frac{1}{N} \sum_{i=1}^N \delta(\lambda - \lambda_i)$, converge vers une mesure de probabilité déterministe μ , i.e. pour toute fonction f continue bornée

$$\frac{1}{N} \sum_{i=1}^N f(\lambda_i) \rightarrow \int f(\lambda) d\mu(\lambda)$$

- On en déduit que

$$\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) \rightarrow \bar{I}_*(\sigma^2) = \int \log \left(1 + \frac{\lambda}{\sigma^2} \right) d\mu(\lambda)$$

$$\text{et que } I(\sigma^2) \rightarrow \bar{I}_*(\sigma^2)$$

- On espère qu'il est plus facile d'étudier l'influence des paramètres de la loi de $\mathbf{H}$ sur $\bar{I}_*(\sigma^2)$

Illustration de la convergence de la distribution des valeurs propres de $\mathbf{H}\mathbf{H}^T$.

Cas où $\mathbf{H} = \frac{1}{\sqrt{n}}\mathbf{X}$, $\mathbf{X}$ iid centrée, variances 1.

$\mathbf{H}\mathbf{H}^T = \frac{1}{n}\mathbf{X}\mathbf{X}^T = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_j \mathbf{x}_j^T$ "matrice de covariance empirique" des vecteurs $(\mathbf{x}_j)_{j=1,\dots,n}$.

Si $\frac{N}{n} \rightarrow 0$, la distribution des valeurs propres converge vers $\delta(\lambda - 1)$.

Si $\frac{N}{n} \rightarrow \alpha$, la distribution converge vers la distribution de Marcenko-Pastur:

- Si $\alpha < 1$, $d\mu(\lambda) = f(\lambda)d\lambda$ avec

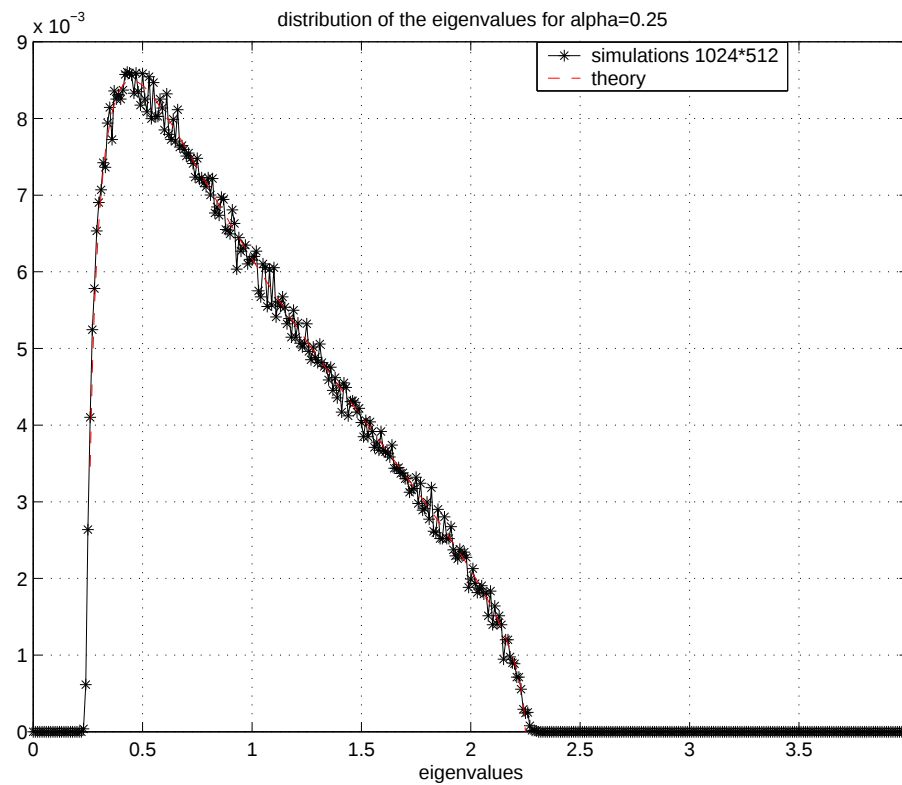
$$f(\lambda) = \frac{1}{2\pi\lambda} \sqrt{[(1 + \sqrt{\alpha})^2 - \lambda][\lambda - (1 - \sqrt{\alpha})^2]} \text{ si } \lambda \in [(\sqrt{\alpha} - 1)^2, (\sqrt{\alpha} + 1)^2]$$

- Si $\alpha > 1$,

$$d\mu(\lambda) = (1 - \alpha)\delta(\lambda) + f(\lambda)d\lambda$$

Illustration dans le cas $n = 1024, N = 256$

Histogrammes des valeurs propres de diverses réalisations de $\mathbf{H}\mathbf{H}^T$.



L'étude de l'information mutuelle moyenne II.

$$I_N(\sigma^2) = \mathbb{E}_{\mathbf{H}} \left[\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) \right]$$

Le schéma général bis.

- $\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right)$ est une fonctionnelle des valeurs propres $(\lambda_i)_{i=1,\dots,N}$ de la matrice de Gram $\mathbf{H}\mathbf{H}^T$ de $\mathbf{H}$:

$$\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) = \frac{1}{N} \sum_{i=1}^N \log \left(1 + \frac{\lambda_i}{\sigma^2} \right)$$

- Sous certaines hypothèses, la distribution empirique des valeurs propres de $\mathbf{H}\mathbf{H}^T$, $\frac{1}{N} \sum_{i=1}^N \delta(\lambda - \lambda_i)$, a le même comportement qu'une mesure de probabilité déterministe μ_N , i.e. pour toute fonction f continue bornée

$$\frac{1}{N} \sum_{i=1}^N f(\lambda_i) - \int f(\lambda) d\mu_N(\lambda) \rightarrow 0$$

- On en déduit que si $\bar{I}_N(\sigma^2) = \int \log(1 + \frac{\lambda}{\sigma^2}) d\mu_N(\lambda)$, alors

$$\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) - \bar{I}_N(\sigma^2) \rightarrow 0$$

$$\text{et que } I_N(\sigma^2) - \bar{I}_N(\sigma^2) \rightarrow 0$$

- On espère qu'il est plus facile d'étudier l'influence des paramètres de la loi de $\mathbf{H}$ sur $\bar{I}_N(\sigma^2)$

L'étude de la distribution des valeurs propres de $\mathbf{H}\mathbf{H}^T$ si $\frac{N}{n} \rightarrow \alpha$

- **Approches combinatoires.**

Méthode des moments: évaluer la limite des moments empiriques $\frac{1}{N} \text{tr}(\mathbf{H}\mathbf{H})^k$, très difficile en général

- **Approches basées sur la transformée de Stieltjes $\frac{1}{N} \left(\sum_{i=1}^N \frac{1}{(\lambda_i - z)} \right)$ de la distribution empirique.**

Montrer que $\frac{1}{N} \left(\sum_{i=1}^N \frac{1}{(\lambda_i - z)} \right) \simeq \int \frac{1}{\lambda - z} d\mu_N(\lambda)$ sur un sous-ensemble suffisamment grand du plan complexe, où μ_N est une mesure de probabilité déterministe.

Approche proposée par Marcenko et Pastur en 1967, puis développée, entre autres, par Girko, Pastur et collègues, et Bai et Silverstein.

L'approche transformée de Stieljès: lemme fondamental (Bai-Silverstein).

Lemme 1 Soient $\mathbf{z} = (z_1, \dots, z_N)^T$ des variables aléatoires i.i.d., centrées, de variance 1, telles que $\mathbb{E}(z_1^{4+\beta}) < +\infty$, et $\mathbf{B}_N$ une matrice aléatoire $N \times N$, indépendante de $\mathbf{z}$ vérifiant $\sup_N \|\mathbf{B}_N\| < K < +\infty$. Alors, si $\frac{N}{n} \rightarrow \alpha$,

$$\mathbb{E} \left| \frac{1}{n} \mathbf{z}^T \mathbf{B}_N \mathbf{z} - \frac{1}{n} \text{tr}(\mathbf{B}_N) \right|^{2+\beta/2} < \frac{C}{n^{1+\beta/4}}$$

où C est une constante ne dépendant que de K et de α .

Implique que $\frac{1}{n} \mathbf{z}^T \mathbf{B}_N \mathbf{z} - \frac{1}{n} \text{tr}(\mathbf{B}_N)$ converge p.s. vers 0.

L'exemple du modèle de Kronecker: préparation.

$\mathbf{H} = \frac{1}{\sqrt{n}} \mathbf{R}^{1/2} \mathbf{X} \tilde{\mathbf{R}}^{1/2}$, $\mathbf{X}$ $N \times n$ iid gaussienne centrée, variances 1, $\mathbf{R} \geq 0$, $\tilde{\mathbf{R}} \geq 0$.

$\mathbf{D}$ et $\tilde{\mathbf{D}}$ matrices diagonales des valeurs propres de $\mathbf{R}$ et $\tilde{\mathbf{R}}$.

On se ramène au cas où $\mathbf{H} = \frac{1}{\sqrt{n}} \mathbf{D}^{1/2} \mathbf{X} \tilde{\mathbf{D}}^{1/2}$.

Les entrées de $\mathbf{H}$ sont indépendantes, mais de variances différentes :

$$\mathbb{E}(H_{i,j}^2) = \frac{1}{n} d_i \tilde{d}_j$$

$\mathbf{h}_j$ colonne j de $\mathbf{H}$ donnée par $\mathbf{h}_j = \frac{1}{\sqrt{n}} \tilde{d}_j^{1/2} \mathbf{D}^{1/2} \mathbf{x}_j$

$\mathbf{h}^{(i)}$ ligne i de $\mathbf{H}$ donnée par $\mathbf{h}^{(i)} = \frac{1}{\sqrt{n}} \mathbf{x}^{(i)} d_i^{1/2} \tilde{\mathbf{D}}^{1/2}$

Les équations fondamentales I.

Quelques notations.

$$\mathbf{Q}(z) = (\mathbf{H}\mathbf{H}^T - z\mathbf{I})^{-1}, \quad \tilde{\mathbf{Q}}(z) = (\mathbf{H}^T\mathbf{H} - z\mathbf{I})^{-1}$$

$\mathbf{h}_j$ colonne j de $\mathbf{H}$, $\mathbf{H}_j$ matrice $\mathbf{H}$ privée de $\mathbf{h}_j$, $\mathbf{Q}_j(z)$ associée à $\mathbf{H}_j$

$\mathbf{h}^{(i)}$ ligne i de $\mathbf{H}$, $\mathbf{H}^{(i)}$ matrice $\mathbf{H}$ privée de $\mathbf{h}^{(i)}$, $\tilde{\mathbf{Q}}^{(i)}(z)$ associée à $\mathbf{H}^{(i)}$

Propriété importante : $\|\mathbf{Q}(z)\| \leq \frac{1}{|\operatorname{Im} z|}, \quad \|\tilde{\mathbf{Q}}(z)\| \leq \frac{1}{|\operatorname{Im} z|}$

Formules de Cramer.

$$q_{ii}(z) = \frac{1}{-z(1 + \mathbf{h}^{(i)}\mathbf{Q}^{(i)}(z)\mathbf{h}^{(i)T})}$$

$$\tilde{q}_{jj}(z) = \frac{1}{-z(1 + \mathbf{h}_j^T\mathbf{Q}_j(z)\mathbf{h}_j)}$$

Les équations fondamentales II.

Etude du comportement de $\mathbf{h}_j^T \mathbf{Q}_j(z) \mathbf{h}_j$

$$\mathbf{h}_j^T \mathbf{Q}_j(z) \mathbf{h}_j = \tilde{d}_j \frac{1}{n} \mathbf{x}_j^T \mathbf{D}^{1/2} \mathbf{Q}_j(z) \mathbf{D}^{1/2} \mathbf{x}_j$$

$$\mathbf{h}_j^T \mathbf{Q}_j(z) \mathbf{h}_j = \tilde{d}_j \frac{1}{n} \text{tr}(\mathbf{D} \mathbf{Q}_j(z)) + \epsilon_{1,j}$$

$$\mathbf{h}_j^T \mathbf{Q}_j(z) \mathbf{h}_j = \tilde{d}_j \frac{1}{n} \text{tr}(\mathbf{D} \mathbf{Q}(z)) + \epsilon_{2,j}$$

Les équations fondamentales III.

$$q_{ii}(z) = \frac{1}{-z(1 + d_i \frac{1}{n} \text{tr}(\tilde{\mathbf{D}}\tilde{\mathbf{Q}}(z)) + \tilde{\epsilon}_{2,i})} = \frac{1}{-z(1 + d_i \frac{1}{n} \text{tr}(\tilde{\mathbf{D}}\tilde{\mathbf{Q}}(z)))} + \tilde{\epsilon}_i$$

$$\tilde{q}_{jj}(z) = \frac{1}{-z(1 + \tilde{d}_j \frac{1}{n} \text{tr}(\mathbf{D}\mathbf{Q}(z)) + \epsilon_{2,j})} = \frac{1}{-z(1 + \tilde{d}_j \frac{1}{n} \text{tr}(\mathbf{D}\mathbf{Q}(z)))} + \epsilon_j$$

Les équations fondamentales IV.

$$\begin{aligned} -z \frac{1}{n} \text{tr}(\mathbf{DQ}(z)) &= \frac{1}{n} \text{tr} \left[\mathbf{D} \left(I + \frac{1}{n} \text{tr}(\tilde{\mathbf{D}}\tilde{\mathbf{Q}}(z)) \mathbf{D} \right)^{-1} \right] + \epsilon \\ -z \frac{1}{n} \text{tr}(\tilde{\mathbf{D}}\tilde{\mathbf{Q}}(z)) &= \frac{1}{n} \text{tr} \left[\tilde{\mathbf{D}} \left(I + \frac{1}{n} \text{tr}(\mathbf{DQ}(z)) \tilde{\mathbf{D}} \right)^{-1} \right] + \tilde{\epsilon} \end{aligned}$$

Les équations fondamentales V.

Soient $\delta(z)$ et $\tilde{\delta}(z)$ les solutions du système.

$$\begin{aligned} -z\delta(z) &= \frac{1}{n} \text{tr} \left[\mathbf{D} \left(I + \tilde{\delta}(z) \mathbf{D} \right)^{-1} \right] \\ -z\tilde{\delta}(z) &= \frac{1}{n} \text{tr} \left[\tilde{\mathbf{D}} \left(I + \delta(z) \tilde{\mathbf{D}} \right)^{-1} \right] \end{aligned}$$

Alors, presque surement, sur $\mathbb{C} - \mathbb{R}^+$,

$$\frac{1}{n} \text{tr}(\mathbf{D} \mathbf{Q}(z)) - \delta(z) \rightarrow 0, \quad \frac{1}{n} \text{tr}(\tilde{\mathbf{D}} \tilde{\mathbf{Q}}(z)) - \tilde{\delta}(z) \rightarrow 0$$

Soit $t_N(z)$ la fonction $t_N(z) = \frac{1}{N} \text{tr} \left[-z \left(\mathbf{I} + \delta(z) \mathbf{D} \right)^{-1} \right]$.

$t_N(z)$ se met sous la forme $t_N(z) = \int_{\mathbb{R}^+} \frac{1}{\lambda - z} d\mu_N(\lambda)$, μ_N est une mesure de probabilité.

Presque surement, sur $\mathbb{C} - \mathbb{R}^+$, $\frac{1}{N} \text{tr}(\mathbf{Q}(z)) - t_N(z) \rightarrow 0$.

Evaluation de l'information mutuelle.

$$\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) = \int_{\sigma^2}^{+\infty} \left(\frac{1}{\omega^2} - \frac{1}{N} \text{Trace} \mathbf{Q}(-\omega^2) \right) d\omega^2$$

$$\frac{1}{N} \log \det \left(\mathbf{I} + \frac{\mathbf{H}\mathbf{H}^T}{\sigma^2} \right) - \int_{\sigma^2}^{+\infty} \left(\frac{1}{\omega^2} - t(-\omega^2) \right) d\omega^2 \rightarrow 0$$

$$\bar{I}_N(\sigma^2) = \int_{\sigma^2}^{+\infty} \left(\frac{1}{\omega^2} - t(-\omega^2) \right) d\omega^2 \text{ vaut aussi}$$

$$\frac{1}{N} \log \det \left(\mathbf{I} + \tilde{\delta}(-\sigma^2) \mathbf{R} \right) + \frac{1}{N} \log \det \left(\mathbf{I} + \delta(-\sigma^2) \tilde{\mathbf{R}} \right) - \sigma^2 \frac{n}{N} \delta(-\sigma^2) \tilde{\delta}(-\sigma^2)$$

$$I_N(\sigma^2) - \bar{I}_N(\sigma^2) \rightarrow 0$$

Le cas non centré I

Le modèle de Kronecker

$$\mathbf{H} = \mathbf{A} + \frac{1}{\sqrt{n}} \mathbf{D}^{1/2} \mathbf{X} \tilde{\mathbf{D}}^{1/2}$$

Dans le cas $\mathbf{A}$ non diagonale, utilisation de la formule de Cramer beaucoup plus difficile : Girko si normes $\mathbb{L}^1$ des lignes et des colonnes de $\mathbf{A}$ sont uniformément bornées. Hypothèse non vérifiée dans les systèmes MIMO.

Approche différente valide dans le cas où les normes $\mathbb{L}^2$ des lignes et des colonnes de $\mathbf{A}$ sont uniformément bornées

$$\frac{1}{N} \text{tr} \mathbf{Q}(-\sigma^2) - \frac{1}{N} \text{tr}(\mathbf{T}(-\sigma^2)) \rightarrow 0$$

Le cas non centré II

$\frac{1}{N} \text{Trace}(\mathbf{T}(-\sigma^2))$ est donné par:

$$\frac{1}{N} \text{Trace}(\mathbf{T}(-\sigma^2)) = \frac{1}{N} \text{Trace} \left[\sigma^2 (\mathbf{I} + \mathbf{D} \tilde{\delta}) + \mathbf{A} (I + \tilde{\mathbf{D}} \delta)^{-1} \mathbf{A}^H \right]^{-1}$$

avec

$$\begin{aligned} \delta &= \frac{1}{N} \text{tr} \left[\mathbf{D} \left(\sigma^2 (\mathbf{I} + \mathbf{D} \tilde{\delta}) + \mathbf{A} (I + \tilde{\mathbf{D}} \delta)^{-1} \mathbf{A}^H \right)^{-1} \right]^{-1} \\ \tilde{\delta} &= \frac{1}{N} \text{tr} \left[\tilde{\mathbf{D}} \left(\sigma^2 (\mathbf{I} + \tilde{\mathbf{D}} \delta) + \mathbf{A}^H (I + \mathbf{D} \tilde{\delta})^{-1} \mathbf{A} \right)^{-1} \right]^{-1} \end{aligned}$$

Le cas non centré III

Approximation de l'information mutuelle.

$$I_N(\sigma^2) - \bar{I}_N(\sigma^2) \rightarrow 0$$

où $\bar{I}_N(\sigma^2)$ est donnée par

$$\frac{1}{N} \log \det \left(\mathbf{I} + \tilde{\delta} \mathbf{D} + \frac{1}{\sigma^2} \mathbf{A} (\mathbf{I} + \tilde{\mathbf{D}} \delta)^{-1} \mathbf{A}^T \right) + \frac{1}{N} \log \det \left(\mathbf{I} + \delta \tilde{\mathbf{D}} \right) - \sigma^2 \frac{n}{N} \delta \tilde{\delta}$$

Qualité de l'approximant déterministe.

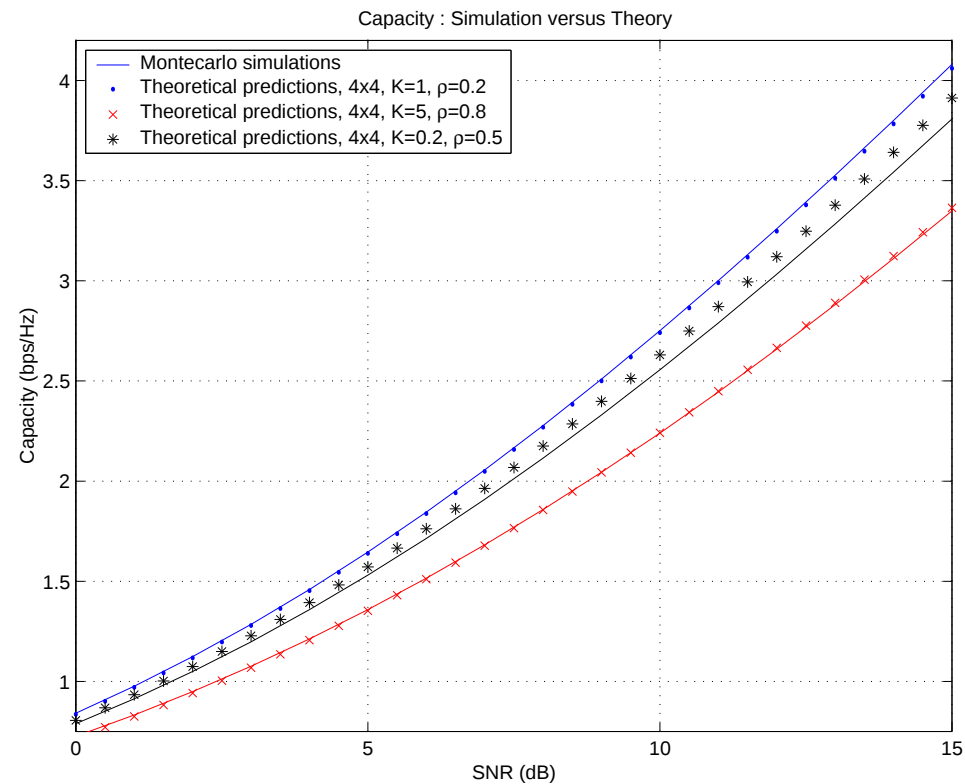


Figure 1: Qualité de l'approximant déterministe.

Conclusions.

- Mise en évidence d'approximations relativement simples (à interpréter et à calculer) de l'information mutuelle de canaux MIMO non centrés
- Bons approximants pour des nombres d'antennes modérés, à confirmer par évaluation des vitesses de convergence
- Possible d'analyser d'autres indicateurs de performances avec cette approche