

Valeurs propres de grandes matrices aléatoires de Gram.
Applications aux systèmes MIMO.

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

Plan de l'exposé.

- Motivations applicatives
- Énoncé des résultats
- Éléments de preuve

Article pouvant être téléchargé sur le site de l'ACI NIM MALCOM,
<http://www.tsi.enst.fr/najim/aci-malcom/>

Les systèmes de communication MIMO.

n antennes d'émissions, N antennes de réceptions.

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

- \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, $\mathbb{E}(\mathbf{v}\mathbf{v}^T) = \sigma^2 I$
- Σ matrice $N \times n$ des gains entre émetteurs et récepteurs

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

La problématique.

Σ réalisation d'un certain type de matrice aléatoire

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

Exemple d'indice de performance usuel.

- Information mutuelle moyenne: $\mathbb{E}_{\Sigma} \left[\frac{1}{N} \log \det \left(I + \frac{\Sigma \Sigma^H}{\sigma^2} \right) \right]$

Débit maximum atteignable par le système MIMO.

- Capacité ergodique: $\text{Max}_{P \geq 0, \frac{1}{n} \text{tr}(P) \leq 1} \mathbb{E}_{\Sigma} \left[\frac{1}{N} \log \det \left(I + \frac{\Sigma P \Sigma^H}{\sigma^2} \right) \right]$

Débit maximum lorsque l'on optimise l'émetteur : remplacer \mathbf{s} par $P_*^{1/2} \mathbf{s}$

Types d'hypothèses sur Σ .

Σ matrice aléatoire gaussienne complexe dans le contexte MIMO (mais non gaussienne dans d'autres contextes)

Différents scénarios considérés dans la littérature MIMO.

- $\Sigma = \frac{1}{\sqrt{n}}X$, X $N \times n$ iid centrée, variances 1.
- $\Sigma = \frac{1}{\sqrt{n}}Z$ avec $Z = R^{1/2}X\tilde{R}^{1/2}$, X $N \times n$ iid centrée, variances 1, $R \geq 0$, $\tilde{R} \geq 0$
Toeplitz, modèle de Kronecker : $\mathbb{E}(Z_{k,l}Z_{p,q}^*) = R_{k,p}\tilde{R}_{l,q}$
- $\Sigma = \frac{1}{\sqrt{n}}Z$, $Z_{k,l}$ 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.

Expression exacte disponible de l'information mutuelle moyenne, mais très peu exploitable.

Etude de l'information mutuelle moyenne en grande dimension.

- Etudier le comportement de l'information mutuelle moyenne 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: cas centré.

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

Le schéma général.

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

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

- Sous certaines hypothèses, la distribution empirique des valeurs propres de $\Sigma \Sigma^H$, $\frac{1}{N} \sum_{i=1}^N \delta(\lambda - \lambda_i)$, 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

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

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

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

- Difficile de calculer "directement" $\bar{I}_*(\sigma^2)$: on passe par l'intermédiaire de $Q(z) = (\Sigma\Sigma^H - zI)^{-1}$ et de $\frac{1}{N} \text{Tr}(Q(z))$.

$$\frac{1}{N} \log \det \left(I + \frac{\Sigma\Sigma^H}{\sigma^2} \right) = \int_{\sigma^2}^{+\infty} \left[\frac{1}{\omega^2} - \frac{1}{N} \text{Tr}(Q(-\omega^2)) \right] d\omega^2$$

$$\frac{1}{N} \text{Tr}(Q(z)) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_i - z} \rightarrow t(z) = \int_{\mathbb{R}^+} \frac{1}{\lambda - z} d\mu(\lambda)$$

$t(z)$ se calcule plus facilement.

$$\bar{I}_*(\sigma^2) = \int_{\sigma^2}^{+\infty} \left[\frac{1}{\omega^2} - t(-\omega^2) \right] d\omega^2$$

Cf. par exemple travaux de Girko (cas non gaussien), Boutet de Monvel-Khorunzhy-Vasilchuck (cas gaussien)

L'étude de l'information mutuelle moyenne: le cas non centré

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

Le schéma général bis.

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

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

- Sous certaines hypothèses, la distribution empirique des valeurs propres de $\Sigma \Sigma^H$, $\frac{1}{N} \sum_{i=1}^N \delta(\lambda - \lambda_i)$, 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 $\bar{I}_n(\sigma^2) = \int \log \left(1 + \frac{\lambda}{\sigma^2} \right) d\mu_n(\lambda)$, alors

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

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

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

- Difficile de calculer "directement" $\bar{I}_n(\sigma^2)$: on passe par l'intermédiaire de $Q(z) = (\Sigma\Sigma^H - zI)^{-1}$ et de $\frac{1}{N}\text{Tr}(Q(z))$.

$$\frac{1}{N} \log \det \left(I + \frac{\Sigma\Sigma^H}{\sigma^2} \right) = \int_{\sigma^2}^{+\infty} \left[\frac{1}{\omega^2} - \frac{1}{N} \text{Tr}(Q(-\omega^2)) \right] d\omega^2$$

$$\frac{1}{N} \text{Tr}(Q(z)) - t_n(z) \rightarrow 0 \text{ avec } t_n(z) = \int_{\mathbb{R}^+} \frac{1}{\lambda - z} d\mu_n(\lambda)$$

$t_n(z)$ se calcule plus facilement.

$$\bar{I}_n(\sigma^2) = \int_{\sigma^2}^{+\infty} \left[\frac{1}{\omega^2} - t_n(-\omega^2) \right] d\omega^2$$

Schéma à utiliser dans le cas de matrices Σ non centrées.

Le cas de canaux non centrés

Le modèle considéré.

$$\Sigma = A + \frac{1}{\sqrt{n}} Z$$

- $Z_{k,l}$ champ stationnaire Gaussien complexe centré
- A matrice déterministe dont les normes des colonnes et des lignes sont uniformément bornées (en n)

Sous certaines hypothèses (cf. exposé de J. Najim), on se ramène à l'étude de

$$\Sigma = A + \frac{1}{\sqrt{n}} Z$$

- $Z_{k,l} = \sigma_{k,l} X_{k,l}$, X matrice iid Gaussienne complexe centrée, éléments de variances 1
- A matrice déterministe dont les normes des colonnes et des lignes sont uniformément bornées (en n)

L'étude de l'information mutuelle moyenne quand $n, N \rightarrow +\infty$, $\frac{N}{n} \rightarrow c$ se ramène à celle de $\frac{1}{N} \text{Tr} [\Sigma \Sigma^H - zI]^{-1}$.

Le problème étudié.

Les hypothèses précises.

$$\Sigma_n = A_n + Y_n$$

- A_n matrice réelle, normes des lignes et les colonnes uniformément bornées
- $Y_{ij}^{(n)} = \frac{1}{\sqrt{n}} \sigma_{ij}^{(n)} X_{ij}^{(n)}$
- Les éléments de X_n sont i.i.d. Moyenne nulle, variance 1, moment d'ordre $4 + \epsilon$ fini
- $\sup_{n,i,j} \sigma_{ij}^{(n)} \leq \sigma_{max} < +\infty$

Travaux précédents dans le cas non centré.

- Dozier-Silverstein (2004), $\sigma_{ij}^{(n)} = 1$, distribution des valeurs propres de $A_n A_n^T$ convergente: la distribution empirique des valeurs propres de $\Sigma_n \Sigma_n^T$ converge presque sûrement vers une mesure de probabilité déterministe.
- Girko (2001) : si $\sup_{n,i} \sum_j |A_{ij}^{(n)}| < +\infty$, et $\sup_{n,j} \sum_i |A_{ij}^{(n)}| < +\infty$, il existe une fonction déterministe à valeurs matricielles $T_n(z)$ telle que $\frac{1}{N} \text{Tr} [(\Sigma_n \Sigma_n^T - zI)^{-1} - T_n(z)] \rightarrow 0$.

Résultats

$$\mathbf{Q}(z) = (\Sigma \Sigma^T - zI)^{-1}$$

$\mathcal{S}(\mathbb{R}^+)$ ensemble des fonctions analytiques sur $\mathbb{C} - \mathbb{R}^+$ s'écrivant sous la forme $\int_0^{+\infty} \frac{1}{\lambda - z} d\mu(\lambda)$ où μ est une mesure de probabilité.

Soit le système d'équations en $(\psi_i(z))_{i=1, \dots, N}, (\tilde{\psi}_j(z))_{j=1, \dots, n}$:

$$\psi_i(z) = \left[-z \left(1 + \frac{1}{n} \sum_{l=1}^n \sigma^2(i, l) \tilde{T}_{l, l}(z) \right) \right]^{-1}$$

$$\tilde{\psi}_j(z) = \left[-z \left(1 + \frac{1}{n} \sum_{k=1}^N \sigma^2(k, j) T_{k, k}(z) \right) \right]^{-1}$$

où

$$T(z) = \left[\Psi^{-1}(z) - A^T z \tilde{\Psi}(z) A \right]^{-1}$$

$$\tilde{T}(z) = \left[\tilde{\Psi}^{-1}(z) - A z \Psi(z) A^T \right]^{-1}$$

$$\Psi(z) = \text{Diag}(\psi_i(z), i = 1, \dots, N), \tilde{\Psi}(z) = \text{Diag}(\tilde{\psi}_j(z), j = 1, \dots, n)$$

- Le système a une solution unique dans $\mathcal{S}(\mathbb{R}^+)^{n+N}$, et $\frac{1}{N} \text{Tr}(T(z)) \in \mathcal{S}(\mathbb{R}^+)$
- ps, $\frac{1}{N} \text{Trace}(Q(z)) - \frac{1}{N} \text{Trace}(T(z)) \rightarrow 0$ sur $\mathbb{C} - \mathbb{R}^+$.
- Posons $\bar{I}_n(\sigma^2) = \int_{\sigma^2}^{+\infty} \left[\frac{1}{\omega^2} - \frac{1}{N} \text{Tr}(T_n(-\omega^2)) \right] d\omega^2$. Alors,

$$\begin{aligned} \bar{I}_n(\sigma^2) &= \frac{1}{N} \log \det \left[\frac{\Psi(-\sigma^2)^{-1}}{\sigma^2} + A \tilde{\Psi}(-\sigma^2) A^T \right] \\ &\quad + \frac{1}{N} \log \det \frac{\tilde{\Psi}(-\sigma^2)^{-1}}{\sigma^2} - \frac{\sigma^2}{nN} \sum_{i,j} \sigma_{ij}^2 T_{ii}(-\sigma^2) \tilde{T}_{jj}(-\sigma^2) \end{aligned}$$

$$\left[\frac{1}{N} \log \det \left(I + \frac{\Sigma_n \Sigma_n^T}{\sigma^2} \right) \right] - \bar{I}_n(\sigma^2) \rightarrow 0$$

$$\mathbb{E} \left[\frac{1}{N} \log \det \left(I + \frac{\Sigma_n \Sigma_n^T}{\sigma^2} \right) \right] - \bar{I}_n(\sigma^2) \rightarrow 0$$

Exemple: champ de variance séparable, $\sigma_{ij}^2 = d_i \tilde{d}_j$

Correspond au modèle de Kronecker.

On pose $D = \text{Diag}(d_i, i = 1, \dots, N)$, $\tilde{D} = \text{Diag}(\tilde{d}_j, j = 1, \dots, N)$,
 $\delta(z) = \frac{1}{n} \text{Tr}(DT(z))$, $\tilde{\delta}(z) = \frac{1}{n} \text{Tr}(\tilde{D}\tilde{T}(z))$.

Il suffit de déterminer $\delta(z)$ et $\tilde{\delta}(z)$ pour résoudre l'équation car

$$\Psi(z) = \left[-z(I + \tilde{\delta}D)\right]^{-1}, \quad \tilde{\Psi}(z) = \left[-z(I + \delta\tilde{D})\right]^{-1}.$$

$$\begin{cases} \delta(z) = \frac{1}{n} \text{Tr} \left[D \left(-z(I + D\tilde{\delta}) + A(I + \tilde{D}\delta)^{-1} A^T \right)^{-1} \right] \\ \tilde{\delta}(z) = \frac{1}{n} \text{Tr} \left[\tilde{D} \left(-z(I + \tilde{D}\delta) + A^T (I + D\tilde{\delta})^{-1} A \right)^{-1} \right] \end{cases} . \quad (1)$$

Evaluation numérique de la pertinence de l'approximant de $I_n(\sigma^2)$.

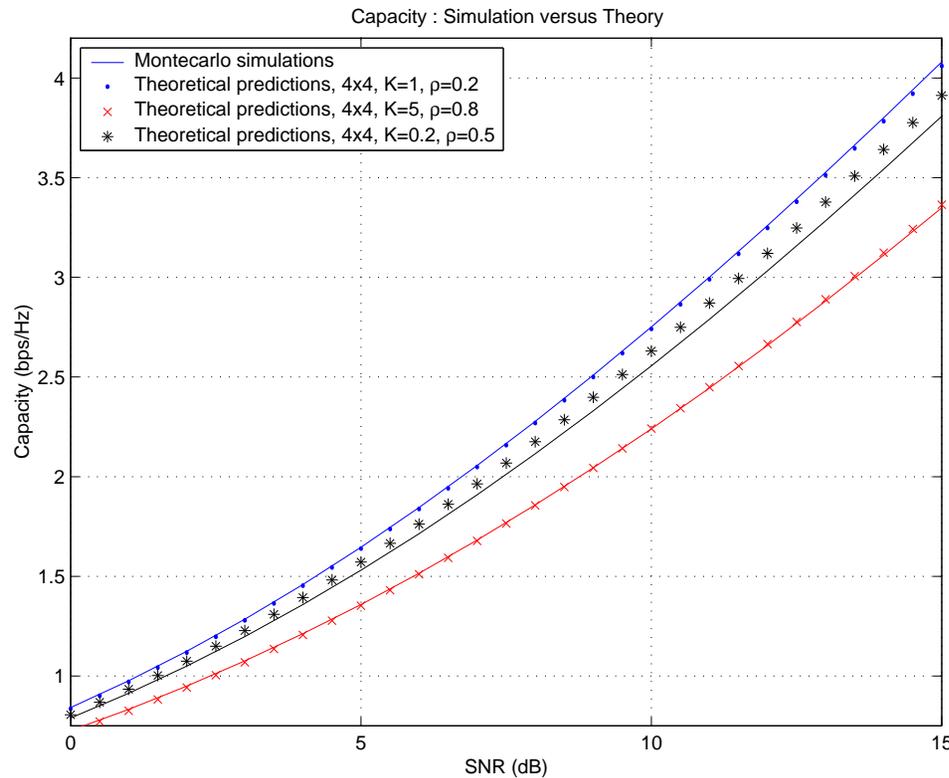


Figure 1: Qualité de l'approximant déterministe, cas Gaussien complexe séparable.

Probable que $I_n(\sigma^2) - \bar{I}_n(\sigma^2) = O(\frac{1}{n^2})$ dans le cas Gaussien complexe séparable.

Etapes de la preuve.

Etape I: montrer l'existence et l'unicité du système d'équations permettant de définir $T(z)$.

Etape II:

Notations.

- $Q(z) = (\Sigma\Sigma^T - zI)^{-1}$, $\tilde{Q}(z) = (\Sigma^T\Sigma - zI)^{-1}$.
- $b_i(z) = \left[-z \left(1 + \frac{1}{n} \sum_{l=1}^n \sigma^2(i, l) \tilde{q}_{l, l}(z)\right)\right]^{-1}$
- $\tilde{b}_j(z) = \left[-z \left(1 + \frac{1}{n} \sum_{k=1}^N \sigma^2(k, j) q_{k, k}(z)\right)\right]^{-1}$
- $R(z) = \left[B^{-1}(z) - zA\tilde{B}(z)A^T\right]^{-1}$
- $\tilde{R}(z) = \left[\tilde{B}^{-1}(z) - zA^TB(z)A\right]^{-1}$

Etape II.

Montrer que pour toutes matrices diagonales U et \tilde{U} tq.
 $\sup_n \|U\| \leq K_1, \sup_n \|\tilde{U}\| \leq \tilde{K}_1$

$$\mathbb{E} \left| \frac{1}{N} \text{Trace} [Q(z)U - R(z)U] \right|^{2+\epsilon/2} \leq \frac{K_2}{n^{1+\epsilon/4}}$$

$$\mathbb{E} \left| \frac{1}{N} \text{Trace} [\tilde{Q}(z)\tilde{U} - \tilde{R}(z)\tilde{U}] \right|^{2+\epsilon/2} \leq \frac{\tilde{K}_2}{n^{1+\epsilon/4}}$$

En prenant $\tilde{U} = \tilde{D}_i = \text{Diag}(\sigma_{i1}, \dots, \sigma_{in}), U = D_j = \text{Diag}(\sigma_{1j}, \dots, \sigma_{Nj})$, en déduire que

$$\frac{1}{n} \sum_{l=1}^n \sigma^2(i, l) \tilde{q}_{l,l}(z) = \frac{1}{n} \sum_{l=1}^n \sigma^2(i, l) \tilde{R}_{l,l}(z) + \epsilon_{i,1}(z)$$

$$\frac{1}{n} \sum_{k=1}^N \sigma^2(k, j) q_{k,k}(z) = \frac{1}{n} \sum_{k=1}^N \sigma^2(k, j) R_{k,k}(z) + \tilde{\epsilon}_{j,1}(z)$$

Etape III: rapprocher $R(z)$ de $T(z)$.

$$b_i(z) = \left[-z \left(1 + \frac{1}{N} \sum_{l=1}^n \sigma^2(i, l) \tilde{R}_{l,l}(z) \right) \right]^{-1} + \epsilon_{i,2}(z)$$

$$\tilde{b}_j(z) = \left[-z \left(1 + \frac{1}{N} \sum_{k=1}^N \sigma^2(k, j) R_{k,k}(z) \right) \right]^{-1} + \tilde{\epsilon}_{j,2}(z)$$

et qui permet de rapprocher $(b_i(z), \tilde{b}_j(z))$ de $(\psi_i(z), \tilde{\psi}_j(z))$, et $\mathbf{R}(z)$ de $\mathbf{T}(z)$ pour z bien choisi, puis pour tout $z \in \mathbb{C} - \mathbb{R}^+$ par prolongement analytique.

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 $\mathbb{E}(z_1^{4+\epsilon}) < +\infty$, et \mathbf{B}_N une matrice aléatoire $N \times N$, indépendante de \mathbf{z} vérifiant $\sup_N \|\mathbf{B}_N\| < K < +\infty$. Alors, si $\frac{N}{n} \rightarrow \alpha$,

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

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