Sur la capacité MIMO en grandes dimensions

W. Hachem (CNRS/ENST), Ph. Loubaton (UMLV) et J. Najim (CNRS/ENST)

GDR ISIS

10 janvier 2007

Problème général

Canal radio à antennes multiples MIMO (Multiple Input Multiple Output) représenté par la matrice alatoire $\mathbf{H}$ de taille $N \times n$.

- $\mathbb{E}\mathbf{H}$ non nécessairement nulle.

- Eléments de $\mathbf{H}$ indépendants mais de variances non nécessairement égales,
  ou
  Eléments de $\mathbf{H} - \mathbb{E}\mathbf{H}$ pris d'un champ stationnaire gaussien.

Information mutuelle moyenne de Shannon par antenne de réception :

$$I(\rho) = \frac{1}{N} \mathbb{E} \log\det \left(\mathbf{I}_N + \frac{1}{\rho} \mathbf{H}\mathbf{H}^* \right)$$

et ρ est la variance du bruit additif gaussien.

Comportement de $I(\rho)$ quand $n \to \infty$ et $\frac{N}{n} \to c > 0$?

Plan

Information mutuelle MIMO et matrices aléatoires

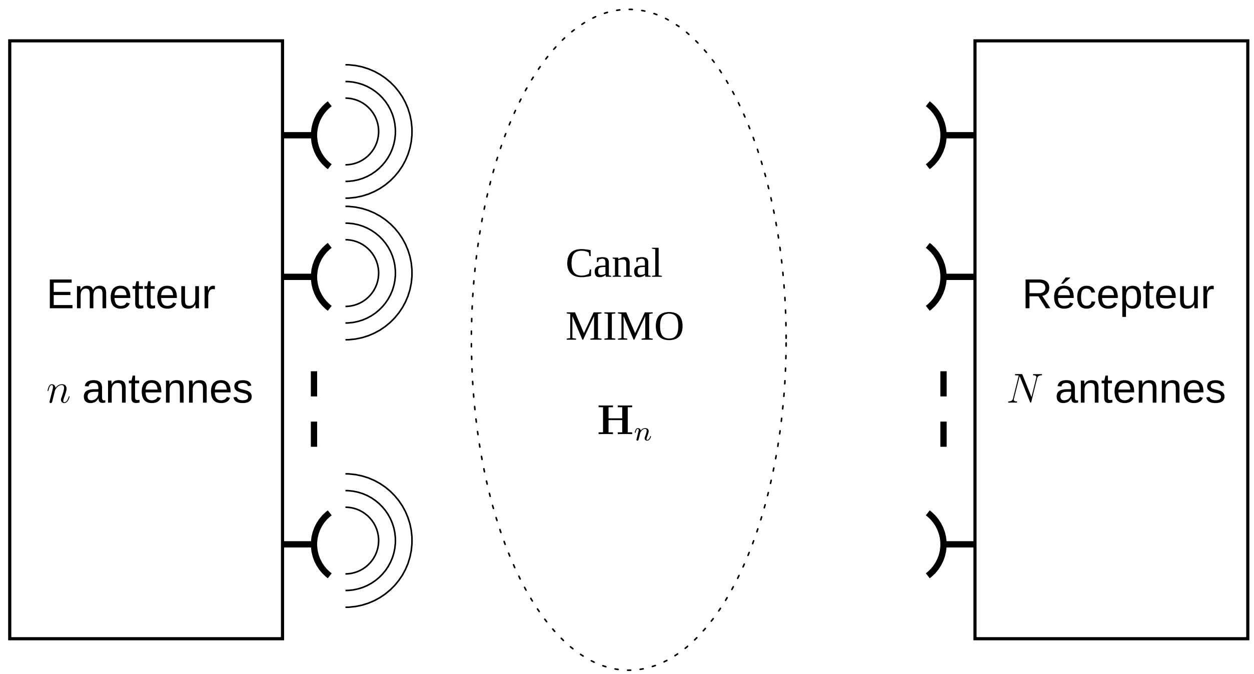

FIG. 1 – Liaison radio à antennes multiples MIMO

Information mutuelle moyenne de Shannon

Information mutuelle moyenne de Shannon par antenne de réception du canal aléatoire $\mathbf{H}_n$ de taille $N \times n$: $I_n(\rho) = \mathbb{E}\mathcal{I}_n(\rho)$ où

$$\mathcal{I}_n(\rho) = \frac{1}{N} \log\det\left(\mathbf{I}_N + \frac{1}{\rho}\mathbf{H}_n\mathbf{H}_n^*\right)$$

où ρ est la variance du bruit additif gaussien.

- **Comportement de $I_n(\rho)$ quand $n \to \infty$ and $\frac{N}{n} \to c > 0$?**
- **Matrice de covariance du signal émis non triviale : optimisation de $I_n(\rho)$?**
- **Fluctuations de $\mathcal{I}_n(\rho)$?**

Information mutuelle et mesure spectrale

- Information mutuelle : $I_n(\rho) = \mathbb{E}\frac{1}{N}\sum_{i=1}^{N}\log\left(1+\frac{\lambda_{i,n}}{\rho}\right)$

  où $\{\lambda_{1,n},\ldots,\lambda_{N,n}\}$ sont les valeurs propres de $\mathbf{H}_n\mathbf{H}_n^*$.

- Mesure spectrale de $\mathbf{H}_n\mathbf{H}_n^*$: mesure de probabilité aléatoire $\mu_n = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_{i,n}}$.

- Donc $I_n(\rho) = \mathbb{E}\int\log\left(1+\frac{t}{\rho}\right)\mu_n(dt)$.

- Etant donné un modèle statistique pour $\mathbf{H}_n$, on espère habituellement que la mesure spectrale μ_n converge faiblement vers une mesure de probabilité déterministe μ quand $n\to\infty$ et $N/n\to c>0$, de manière à avoir

$$I_n(\rho) \xrightarrow[n\to\infty]{} I^*(\rho) = \int\log\left(1+\frac{t}{\rho}\right)\mu(dt) \; .$$

  μ : Mesure Spectrale Limite (MSL).

LE CAS OÙ $\mathbf{H}$ EST À ÉLÉMENTS INDEPENDANTS $\mathcal{CN}(0,1/n)$

L'exemple classique (Telatar'98) : éléments de $\mathbf{H}_n$ complexes Gaussiens indépendants centrés de variance $1/n$.

Marchenko et Pastur'67 : μ_n converge faiblement vers la MSL

$$\mu_c(dt) = \max\left(1 - \frac{1}{c}, 0\right) \delta_0(t) + f(t)dt$$

où $f(t) = \dfrac{\sqrt{(b_+ - t)(t - b_-)}}{2\pi c t} \mathbf{1}_{[b_-, b_+]}(t)$ and $b_\pm = \left(\sqrt{c} \pm 1\right)^2$.

Note : caractère gaussien non nécessaire pour montrer cette convergence.

Loi de Marchenko-Pastur

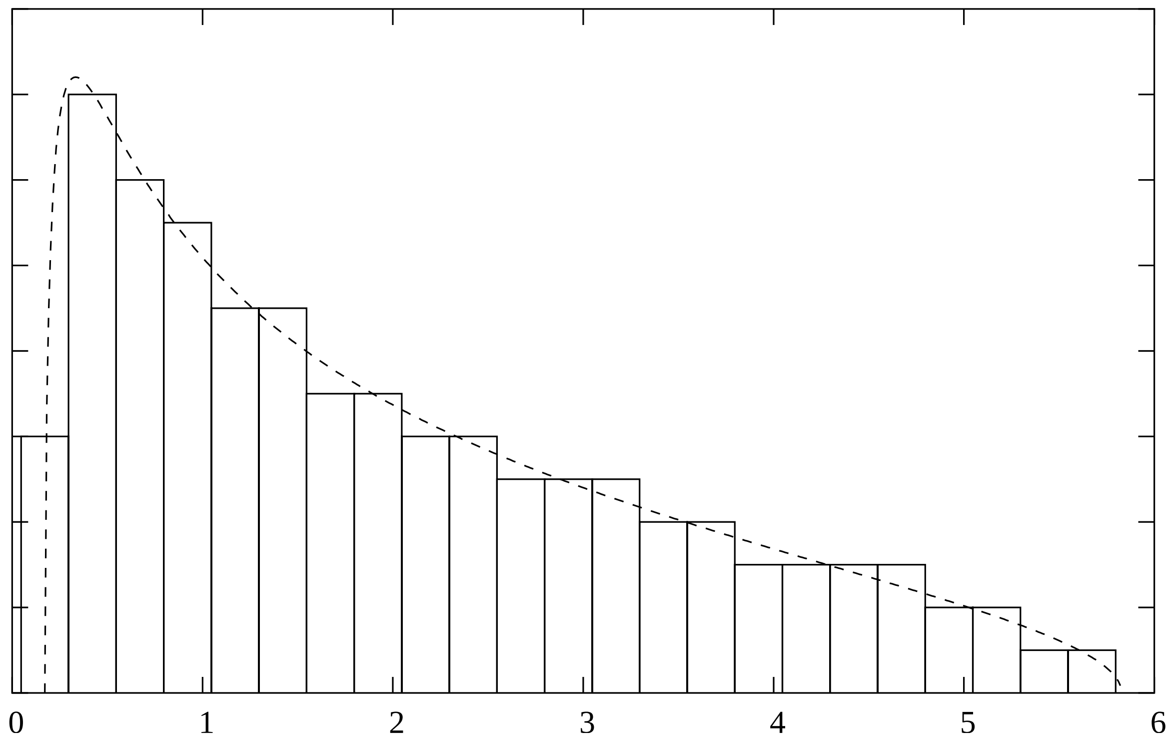

Fig. 2 – Densité de Marchenko-Pastur pour $c = 1/2$

La transformée de Stieltjes

- Transformée de Stieltjes (TS) $m_\mu(z)$ d'une mesure de probabilité μ : la fonction complexe
$$m_\mu(z) = \int \frac{1}{t-z}\mu(dt)$$
analytique sur $\mathbb{C} -$ support(μ).

- Convergence faible des mesures de probabilité μ_n vers μ équivaut à la convergence ponctuelle de $m_{\mu_n}(z)$ vers la TS $m_\mu(z)$.

- En théorie des matrices aléatoires, l'existence de la MSL μ est souvent prouvée en prouvant la convergence des $m_{\mu_n}(z)$ vers une TS déterministe $m_\mu(z)$. Fonction $m_\mu(z)$ définie d'habitude comme la solution unique d'une équation fonctionnelle ou d'un système d'équations fonctionnelles.

Présentation du problème

$\boxed{\text{Canal : modèle statistique 1}}$

Canal dit de “Rice” :

$$\mathbf{H}_n = \mathbf{Z}_n + \mathbf{B}_n$$

- $\mathbf{Z}_n = \left[Z_{i,j}^{(n)}\right]$, éléments d’un champ stationnaire gaussien 2D de fonction de covariance γ :

$$\mathbb{E}\left[Z_{i_1,j_1}^{(n)} Z_{i_2,j_2}^{(n)}{}^{*}\right] = \frac{1}{n}\gamma\left(i_1 - i_2, j_1 - j_2\right)$$

- $\mathbf{B}_n$ est une matrice déterministe ("composante de Rice").
  Note : rang de $\mathbf{B}_n$ peut être d’ordre $\mathcal{O}(n)$.

Canal : modèle statistique 2

Matrice du canal : $\mathbf{F}_N \mathbf{H}_n \mathbf{F}_n^*$ où $\mathbf{F}_l$ est la matrice de Fourier $l \times l$ et

$$\mathbf{H}_n = \mathbf{Y}_n + \mathbf{A}_n$$

- $\mathbf{Y}_n = \left[Y_{i,j}^{(n)} \right]$ où $Y_{i,j}^{(n)} = \dfrac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}$, variables aléatoires X_{ij} gaussiennes standard indépendantes.
- $\mathbf{A}_n$ matrice déterministe, rang d'ordre $\mathcal{O}(n)$.

Remarques :

- Un exemple :
  Modèle de Sayeed'02 : le canal est la transformée de Fourier 2D de $\mathbf{Y}_n$; son information mutuelle est $I_n(\rho)$ pour $\mathbf{A}_n = \mathbf{0}$.
- Un cas particulier :
  Hypothèse **(A)** : le profil de variance $\sigma_{ij}(n)^2 = \sigma^2 \left(\dfrac{i}{N}, \dfrac{j}{n} \right)$ où $\sigma^2(x, y)$ est une fonction continue sur $[0, 1]^2$: "profil de variance limite".

LIEN ENTRE LES MODÈLES 1 ET 2

Pour l'étude asymptotique, le modèle 1 peut être remplacé par le modèle 2 avec

- l'hypothèse **(A)** où $\sigma^2(x,y) = \Gamma(x,y)$ where

$$\Gamma(x,y) = \sum_{i,j} \gamma(i,j) e^{-2\imath\pi(ix-jy)}$$

  est la Densité Spectrale de Puissance du processus $Z_{i,j}$.
- $\mathbf{A}_n$ est la transformée de Fourier 2D de $\mathbf{B}_n$.

cf. HLN'05.

Présentation du problème

Modèle 2 : $\mathbf{H}_n = \mathbf{Y}_n + \mathbf{A}_n$ de taille $N \times n$.

- $\mathbf{Y}_n = \left[Y_{i,j}^{(n)} \right]$ où $Y_{i,j}^{(n)} = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}$, variables aléatoires X_{ij} iid centrées de variance 1.
  Les X_{ij} ne sont pas nécessairement gaussiennes.
- $\mathbf{A}_n$ matrice déterministe.

- Etudier le comportement asymptotique de la mesure spectrale μ_n of $\mathbf{H}_n \mathbf{H}_n^*$ quand $n \to \infty$ et $N/n \to c > 0$, où, d'une manière équivalente, sa TS $m_{\mu_n}(z)$.
- En déduire le comportement asymptotique de l'information mutuelle de Shannon $I_n(\rho)$.

Quelques cas particuliers connus

Le cas centré $\mathbb{E}\mathbf{H} = \mathbf{0}$:

- Fonction de covariance séparable $\gamma(i,j) = \gamma_{\mathsf{R}}(i)\gamma_{\mathsf{T}}(j)$ pour le modèle 1 (de manière équivalente, profil de variance séparable $\sigma^2_{ij}(n) = \sigma^2_{\mathsf{R},i}(n)\sigma^2_{\mathsf{T},j}(n)$ pour le modèle 2) : cas étudié par

  Chuah *et.al.*'02 : approche basée sur un résultat de Girko utilisant la TS.

  Moustakas *et.al.*'03 : méthode des répliques.

  Cas séparable : modèle dit de Kronecker dans la littérature MIMO.
- Cas où γ non nécessairement séparable étudié par Tulino *et.al.* : Girko également.

Cas non centré :

- Cas où les éléments de $\mathbf{Y}_n$ sont iid étudié par

  Cottatellucci et Debbah'04 utilisant Girko.

  Moustakas *et.al.* : méthode des répliques.
- Existence d'un profil de variance séparable : Taricco'06 par la méthode des répliques.

Observations préliminaires

Remarques générales sur l'existence d'une MSL

- Dans le cas centré ($\mathbf{A}_n = \mathbf{0}$), une hypothèse sur $\sigma^2_{i,j}$ du type **(A)** *i.e.*, l'existence d'un profil de variance limite, est nécessaire pour établir l'existence d'une MSL.

- Dans le cas non centré ($\mathbf{A}_n \neq \mathbf{0}$) : $\mathbf{H}\mathbf{H}^*$ ne possède de MSL que dans quelques situations particulières. Exemples de telles situations :
  - $\mathbf{Y}_n$ à éléments iid, *i.e.*, profil de variance constant, et $\mathbf{A}_n\mathbf{A}_n^*$ a une MSL (Brent Dozier and Silverstein'04).
  - Profil de variance non nécessairement trivial, **(A)** est satisfaite, mais $\mathbf{A}_n$ est diagonale + une hypothèse supplémentaire sur la diagonale de $\mathbf{A}_n$ (*cf.* HLN, *Annales de l'IHP*'05).

  Dans le cas "général" non centré, $\mathbf{H}\mathbf{H}^*$ n'a pas de MSL.

L'APPROXIMATION DÉTERMINISTE

En revanche, il est possible de produire une approximation déterministe de μ_n : $\exists$ une matrice $N \times N$ déterministe $\mathbf{T}_n(z)$ pour laquelle

$$m_{\mu_n}(z) - \frac{1}{N}\mathrm{tr}\mathbf{T}_n(z) \xrightarrow[n\to\infty]{} 0 \text{ presque sûrement sur } \mathbb{C} - \mathbb{R}_+ \ .$$

$\frac{1}{N}\mathrm{tr}\mathbf{T}_n(z)$ est la TS $m_{\nu_n}(z)$ d'une mesure de probabilité déterministe ν_n qui approxime μ_n dans le sens où pour toute fonction continue bornée $f(t)$,

$$\int f(t)\mu_n(dt) - \int f(t)\nu_n(dt) \xrightarrow[n\to\infty]{} 0$$

Dans les cas particuliers où existe une MSL μ, la TS $m_{\nu_n}(z)$ converge vers $m_\mu(z)$.

TECHNIQUE

- La résolvante de $\mathbf{H}_n\mathbf{H}_n^*$ est $\mathbf{Q}_n(z) = (\mathbf{H}_n\mathbf{H}_n^* - z\mathbf{I}_N)^{-1}$. Nous avons

$$\begin{aligned} \frac{1}{N}\mathrm{tr}\mathbf{Q}_n(z) &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_{i,n} - z} = \frac{1}{N}\sum_{i=1}^{N}\int\frac{1}{t-z}\delta_{\lambda_{i,n}} \\ &= \int\frac{1}{t-z}\mu_n(dt) = m_{\mu_n}(z) \end{aligned}$$

- Technique : considérer les diagonales des résolvantes $\mathbf{Q}_n(z)$ et $\widetilde{\mathbf{Q}}_n(z) = (\mathbf{H}_n^*\mathbf{H}_n - z\mathbf{I}_n)^{-1}$.

- $\frac{1}{N}\mathrm{tr}\mathbf{T}_n(z)$ sera l'approximant de $\frac{1}{N}\mathrm{tr}\mathbf{Q}_n(z)$.

Comportement asymptotique de l'information mutuelle

En dérivant $I_n(\rho)$ par rapport à ρ, nous avons

$$I_n(\rho) = \int_\rho^\infty \left(\frac{1}{\omega} - \mathbb{E}\frac{1}{N}\mathrm{tr}\mathbf{Q}_n(-\omega) \right) d\omega = \int_\rho^\infty \left(\frac{1}{\omega} - \mathbb{E}m_{\mu_n}(-\omega) \right) d\omega$$

$\mathbb{E}m_{\mu_n}(-\omega)$ est approximée par $m_{\nu_n}(z) = \dfrac{1}{N}\mathrm{tr}\mathbf{T}_n(z)$.

Expliciter $\displaystyle\int_\rho^\infty \left(\frac{1}{\omega} - m_{\nu_n}(-\omega) \right) d\omega$

Résultats dans le cas centré

Avec quelques hypothèses (*cf.* le cas général ci-dessous),

- Le système de N équations fonctionnelles

$$t_i(z) = \frac{1}{-z + \frac{1}{n}\sum_{j=1}^{n} \frac{\sigma_{ij}^2}{1 + \frac{1}{n}\sum_{l=1}^{N} \sigma_{lj}^2 t_l(z)}}$$

admet une solution unique $(t_1(z), \cdots, t_N(z))$ où les $t_i(z)$ sont des TS de mesures de probabilité.

- Presque sûrement,

$$\frac{1}{N}\mathrm{tr}\mathbf{Q}_n(z) - \frac{1}{N}\sum_{i=1}^{N} t_i(z) \xrightarrow[n\to\infty]{} 0 \quad \forall z \in \mathbb{C} - \mathbb{R}_+ \ .$$

L'approximant déterministe est la matrice $\mathbf{T}(z) = \mathrm{diag}(t_i(z))$.

Approximation de l'information mutuelle

A partir de $I_n(\rho) = \int_\rho^\infty \left(\frac{1}{\omega} - \mathbb{E} \frac{1}{N} \mathrm{tr} \mathbf{Q}_n(-\omega) \right) d\omega$ et du résultat précédent, nous obtenons :

La quantité

$$\bar{I}_n(\rho) = -\frac{1}{N} \sum_{i=1}^{N} \log \rho t_i(-\rho) + \frac{1}{N} \sum_{j=1}^{n} \log \left(1 + \frac{1}{n} \sum_{l=1}^{N} \sigma_{lj}^2 t_l(-\rho) \right) - \frac{1}{Nn} \sum_{i=1:N, j=1:n} \frac{\sigma_{ij}^2 t_i(-\rho)}{1 + \frac{1}{n} \sum_{l=1}^{N} \sigma_{lj}^2 t_l(-\rho)}$$

satisfait

$$I_n(\rho) - \bar{I}_n(\rho) \xrightarrow[n \to \infty]{} 0.$$

CAS OÙ L'HYPOTHÈSE **(A)** EST SATISFAITE

Ici, nous avons une MSL μ. Sa TS s'écrit

$$m_\mu(z) = \int_0^1 k(u, z) du$$

où $k(u, z)$ est la solution unique de l'équation

$$k(u, z) = \frac{1}{-z + \displaystyle\int_0^1 \frac{\sigma^2(u, t)}{1 + c\int_0^1 \sigma^2(x, t) k(x, z) dx} dt}$$

dans la classe des fonctions $k(u, z)$ telles que

$k(u, z)$ est la TS d'une mesure de probabilité pour tout u.

$k(u, z)$ est continue en u.

Cas où l'hypothèse **(A)** est satisfaite : l'information mutuelle

Grâce à l'existence d'une MSL, $I_n(\rho)$ converge :

$$I_n(\rho) \xrightarrow[n\to\infty]{} I^*(\rho)$$

$$\begin{aligned} I^*(\rho) \quad = \quad & -\int_0^1 \log\left(\rho k\left(u,-\rho\right)\right)\,du \\ & +\frac{1}{c}\int_0^1 \log\left(1+c\int_0^1 \sigma^2(x,u)k\left(x,-\rho\right)dx\right)du \\ & -\int_{[0,1]^2} \frac{\sigma^2(x,y)k\left(x,-\rho\right)}{1+c\int_0^1 \sigma^2(u,y)k\left(u,-\rho\right)du}dx\,dy \end{aligned}$$

Le cas général (non centré)

Théorème 1 *Dans le modèle $\mathbf{H}_n = \mathbf{Y}_n + \mathbf{A}_n$, on suppose que*

- $Y_{i,j}^{(n)} = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}$ *où les X_{ij} sont iid, centrés et de variance 1 et $\mathbb{E}\,|X_{11}|^{4+\varepsilon} < \infty$ pour un $\varepsilon > 0$.*
- $\sup\limits_{i,j,n} \sigma_{ij}^2(n) < \infty$.
- *Normes euclidiennes des lignes et des colonnes de $\mathbf{A}_n$ uniformément bornées.*

Soit $\mathbf{D}_j = \mathrm{diag}\left(\left[\sigma_{1j}^2, \ldots, \sigma_{Nj}^2\right]\right)$ *et* $\widetilde{\mathbf{D}}_i = \mathrm{diag}\left(\left[\sigma_{i1}^2, \ldots, \sigma_{in}^2\right]\right)$.

Le système déterministe de $N+n$ équations :

$$\psi^{(i)}(z) \;=\; \frac{-1}{z\left(1+\frac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{D}}^{(i)}\widetilde{\mathbf{T}}(z)\right)\right)} \quad \text{for } 1 \le i \le N,$$

$$\tilde{\psi}^{(j)}(z) \;=\; \frac{-1}{z\left(1+\frac{1}{n}\mathrm{tr}\left(\mathbf{D}^{(j)}\mathbf{T}(z)\right)\right)} \quad \text{for } 1 \le j \le n,$$

où

$$\mathbf{\Psi}(z) \;=\; \mathrm{diag}\left(\left[\psi^{(1)}(z),\ldots,\psi^{(N)}(z)\right]\right), \quad \widetilde{\mathbf{\Psi}}(z) \;=\; \mathrm{diag}\left(\left[\tilde{\psi}^{(1)}(z),\ldots,\tilde{\psi}^{(n)}(z)\right]\right)$$

$$\mathbf{T}(z) \;=\; \left(\mathbf{\Psi}^{-1}(z) - z\mathbf{A}\widetilde{\mathbf{\Psi}}(z)\boldsymbol{A}^*\right)^{-1}, \quad \widetilde{\mathbf{T}}(z) \;=\; \left(\widetilde{\mathbf{\Psi}}^{-1}(z) - z\mathbf{A}^*\mathbf{\Psi}(z)\mathbf{A}\right)^{-1}$$

admet une solution unique $(\psi^{(1)},\ldots,\psi^{(N)},\tilde{\psi}^{(1)},\ldots,\tilde{\psi}^{(n)})$ *dans la classe des TS de mesures de probabilité portées par* $\mathbb{R}_+$. *Par ailleurs, presque sûrement,*

$$\left(\frac{1}{N}\mathrm{tr}\mathbf{Q}_n(z) - \frac{1}{N}\mathrm{tr}\mathbf{T}_n(z)\right) \xrightarrow[n\to\infty]{} 0 \quad \forall z \in \mathbb{C}-\mathbb{R}_+,$$

$$\left(\frac{1}{n}\mathrm{tr}\widetilde{\mathbf{Q}}_n(z) - \frac{1}{n}\mathrm{tr}\widetilde{\mathbf{T}}_n(z)\right) \xrightarrow[n\to\infty]{} 0 \quad \forall z \in \mathbb{C}-\mathbb{R}_+$$

Commentaires

Extension d'un résultat de Girko et simplification de sa preuve. Approximation de l'information mutuelle de Shannon.

Hypothèses réalistes en communications numériques. En particulier, les colonnes de la "composante de Rice" sont typiquement de la forme

$$\frac{C}{\sqrt{N}} \begin{bmatrix} 1 \\ \exp(\imath\omega) \\ \vdots \\ \exp(\imath(N-1)\omega) \end{bmatrix}$$

en tant que vecteurs de directions d'arrivée. Norme euclidienne bornée.

CAS GÉNÉRAL : APPROXIMATION DE L'INFORMATION MUTUELLE

A partir de $I_n(\rho) = \int_\rho^\infty \left(\frac{1}{\omega} - \mathbb{E}\frac{1}{N}\mathrm{tr}\mathbf{Q}_n(-\omega) \right) d\omega$ et du théorème 1, nous pouvons établir :

Théorème 2 *Soit*

$$\overline{I}_n(\rho) = \frac{1}{N}\log\det\left[\frac{\mathbf{\Psi}(-\rho)^{-1}}{\rho} + \mathbf{A}\widetilde{\mathbf{\Psi}}(-\rho)\mathbf{A}^* \right] + \frac{1}{N}\log\det\frac{\widetilde{\mathbf{\Psi}}(-\rho)^{-1}}{\rho} - \frac{\rho}{Nn}\sum_{\substack{i=1:N\\ j=1:n}} \sigma_{ij}^2 t_i(-\rho)\tilde{t}_j(-\rho)$$

où t_i et $\tilde{t}_j$ sont les éléments diagonaux de $\mathbf{T}_n(z)$ et de $\widetilde{\mathbf{T}}_n(z)$. Alors

$$I_n(\rho) - \overline{I}_n(\rho) \xrightarrow[n\to\infty]{} 0.$$

Le cas non centré séparable

- Modèle de Kronecker ($\sigma^2_{ij}(n) = \sigma^2_{\mathsf{R},i}(n)\sigma^2_{\mathsf{T},j}(n)$) avec une "composante de Rice" ($\mathbf{A}_n \neq \mathbf{0}$).

- Ainsi,

$$\mathbf{H} = \frac{1}{\sqrt{n}}\mathbf{D}_{\mathsf{R}}^{1/2}\mathbf{X}\mathbf{D}_{\mathsf{T}}^{1/2} + \mathbf{A}$$

  où $\mathbf{X}$ est à éléments iid centrés de variance 1 et $\mathbf{D}_{\mathsf{R}} = \mathrm{diag}\left(\sigma^2_{\mathsf{R},1}, \ldots, \sigma^2_{\mathsf{R},N}\right)$ et $\mathbf{D}_{\mathsf{T}} = \mathrm{diag}\left(\sigma^2_{\mathsf{T},1}, \ldots, \sigma^2_{\mathsf{T},n}\right)$.

- Le système de $N + n$ équations qui donne $\mathbf{T}(z)$ et $\widetilde{\mathbf{T}}(z)$ se ramène à un système de deux équations.

APPROXIMANT DÉTERMINISTE

L'approximation déterministe de l'information mutuelle devient

$$\overline{I}_n(\rho) = \frac{1}{N}\log\det\left(\mathbf{I} + \delta\mathbf{D}_{\mathsf{T}} + \frac{1}{\rho}\mathbf{A}^*\left(\mathbf{I} + \tilde{\delta}\mathbf{D}_{\mathsf{R}}\right)^{-1}\mathbf{A}\right) + \frac{1}{N}\log\det\left(\mathbf{I} + \tilde{\delta}\mathbf{D}_{\mathsf{R}}\right) - \rho\frac{n}{N}\delta\tilde{\delta}$$

où $(\delta, \tilde{\delta})$ est la solution unique de

$$\begin{aligned} \delta &= \frac{1}{n}\mathsf{tr}\left(\mathbf{D}_{\mathsf{R}}\left(\rho\left(\mathbf{I} + \tilde{\delta}\mathbf{D}_{\mathsf{R}}\right) + \mathbf{A}\left(\mathbf{I} + \delta\mathbf{D}_{\mathsf{T}}\right)^{-1}\mathbf{A}^*\right)^{-1}\right) \\ \tilde{\delta} &= \frac{1}{n}\mathsf{tr}\left(\mathbf{D}_{\mathsf{T}}\left(\rho\left(\mathbf{I} + \delta\mathbf{D}_{\mathsf{T}}\right) + \mathbf{A}^*\left(\mathbf{I} + \tilde{\delta}\mathbf{D}_{\mathsf{R}}\right)^{-1}\mathbf{A}\right)^{-1}\right) \end{aligned}$$

Application : approximation de la capacité de Shannon

- Problème : soit le canal MIMO de taille $N \times n$

$$\mathbf{H} = \mathbf{\Gamma}_{\mathsf{R}}^{1/2} \mathbf{X} \mathbf{\Gamma}_{\mathsf{T}}^{1/2} + \mathbf{A}$$

  où $\mathbf{A}$ est la "composante de Rice", $\mathbf{X}$ est à éléments gaussiens iid $\sim \mathcal{CN}(0, 1/n)$ et $\mathbf{\Gamma}_{\mathsf{T}}$ et $\mathbf{\Gamma}_{\mathsf{R}}$ sont les matrices de covariance à l'émission et à la réception.
- Trouver une matrice semi définie positive $\mathbf{Q}$ solution de

$$\max_{\mathbf{Q} \geq 0, \mathrm{tr}\mathbf{Q} \leq n} \mathbb{E} \log \det \left(\mathbf{I}_N + \frac{1}{\rho} \mathbf{H} \mathbf{Q} \mathbf{H}^* \right) \quad .$$

- Solution directe très difficile en général.
- La matrice $\mathbf{HQH}^*$ possède la même loi de probabilité qu'une matrice $\mathbf{U \Sigma \Sigma^* U^*}$ où $\mathbf{U}$ est unitaire et $\mathbf{\Sigma}$ correspond au modèle non centré séparable décrit plus haut.
- Il serait alors pertinent de maximiser l'approximation déterministe de $I(\rho)$ par rapport à $\mathbf{Q}$.

Approximation de la capacité

Approximant déterministe de l'information mutuelle moyenne :

Equations du transparent 27 dans lesquelles on remplace
- $\mathbf{A}$ par $\mathbf{A}\mathbf{Q}^{1/2}$,
- $\mathbf{D}_{\mathsf{T}}$ par $\mathbf{Q}^{1/2}\mathbf{\Gamma}_{\mathsf{T}}\mathbf{Q}^{1/2}$,
- $\mathbf{D}_{\mathsf{R}}$ par $\mathbf{\Gamma}_{\mathsf{R}}$.

Ces équations prennent la forme

$$\overline{I}(\rho) = \log\det\left(\mathbf{I} + \mathbf{G}\left(\delta, \tilde{\delta}\right)\mathbf{Q}\right) + h(\delta, \tilde{\delta})$$

où

$$\mathbf{G}\left(\delta, \tilde{\delta}\right) = \delta\mathbf{\Gamma}_{\mathsf{T}} + \frac{1}{\rho}\mathbf{A}^*\left(\mathbf{I} + \tilde{\delta}\mathbf{\Gamma}_{\mathsf{R}}\right)^{-1}\mathbf{A}$$

et $(\delta, \tilde{\delta})$ est la solution d'un système qui dépend de $\mathbf{Q}$:

$$\begin{aligned} \delta &= f\left(\delta, \tilde{\delta}, \mathbf{Q}\right) \\ \tilde{\delta} &= \tilde{f}\left(\delta, \tilde{\delta}, \mathbf{Q}\right) \end{aligned}$$

MATRICE DE COVARIANCE $\mathbf{Q}$ QUI ATTEINT LA CAPACITÉ

Pour un couple $(\delta, \tilde{\delta})$ donné, la matrice $\mathbf{Q}$ qui rend $\log\det\left(\mathbf{I} + \mathbf{G}(\delta, \tilde{\delta})\mathbf{Q}\right)$ maximum est obtenue par "water filling".

Algorithme :

1. Poser $\mathbf{Q}_0 = \mathbf{I}_n$, résoudre le système d'équations $\delta = f\left(\delta, \tilde{\delta}, \mathbf{Q}_0\right)$ et $\tilde{\delta} = \tilde{f}\left(\delta, \tilde{\delta}, \mathbf{Q}_0\right)$ pour obtenir δ_0 et $\tilde{\delta}_0$.
2. Calculer $\mathbf{Q}_1$ en rendant $\log\det\left(\mathbf{I} + \mathbf{G}(\delta_0, \tilde{\delta}_0)\mathbf{Q}\right)$ maximum par "water filling".
3. Itérer.

Convergence, Théorème de la Limite Centrale (TLC)

Rappelons que $\mathcal{I}_n(\rho) = \dfrac{1}{N} \log\det\left(\mathbf{I}_N + \dfrac{1}{\rho}\mathbf{H}_n\mathbf{H}_n^*\right)$ et que $I_n(\rho) = \mathbb{E}\mathcal{I}_n(\rho)$.

- TLC sur $\mathcal{I}_n$ pour $n \to \infty$ et $N/n \to c > 0$, au moins dans le cas centré $\mathbf{A}_n = \mathbf{0}$ (que nous supposerons).
- Par l'intermédiaire de "l'approximation gaussienne", nous nous faisons une idée de la "probabilité de coupure" $\mathbb{P}(\mathcal{I}_n < \text{un seuil } R)$.
- Deux termes :
  - TLC sur la variable aléatoire $\chi_{1,n} = N(\mathcal{I}_n - I_n)$, en particulier, approximation de la variance de $\chi_{1,n}$.
  - Biais $\chi_{2,n} = N\left(I_n - \overline{I}_n\right)$ entre l'information mutuelle NI_n et l'approximation déterministe $N\overline{I}_n$.

TLC : RÉSULTATS GÉNÉRAUX

Technique : TLC pour les martingales comme dans Girko (méthode REFORM) et dans Bai et Silverstein'04.
Dans le cas où les éléments de $\mathbf{Y}_n$ sont gaussiens, on peut aussi utiliser une technique développée par Pastur et ses collègues.

- Il existe une suite déterministe $\theta_n = \mathcal{O}(1)$ telle que $\chi_{1,n}/\theta_n$ converge en loi vers $\mathcal{N}(0, 1)$.
  $\Rightarrow$ La variance de $\mathcal{I}_n$ décroît en $\mathcal{O}(1/N^2)$.
- Biais : $\chi_{2,n} \approx \left(\mathbb{E}\,|X_{11}|^4 - 2\right) \times \mathsf{Cst}_n + o(1)$.
  $\Rightarrow$ Négligeable si les éléments de $\mathbf{Y}_n$ sont gaussiens (canal "de Rayleigh").