Spectral Measure of Certain Gram Random Matrices
Applications in Wireless Communications

Walid Hachem (Supélec, France)

Work done with Philippe Loubaton (UMLV) and Jamal Najim (CNRS)

NUS / IMS Program on Random Matrices

23 March 2006

Gram matrices in this presentation

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

- $\mathbf{Y}_n$ is a $N \times n$ random matrix with independent centered elements having possibly different variances.
- $\mathbf{A}_n$ is a deterministic matrix.

Eigenvalue distribution of $\mathrm{H}_n \mathrm{H}_n^{\mathrm{H}}$ when $n \to \infty$ and $\dfrac{N}{n} \to c > 0$?

Outline

1) **Problem statement**

2) **Some particular cases**

3) **The general case**

4) **The general case: main steps of the proof**

5) **Towards a Central Limit Theorem**

Problem Statement

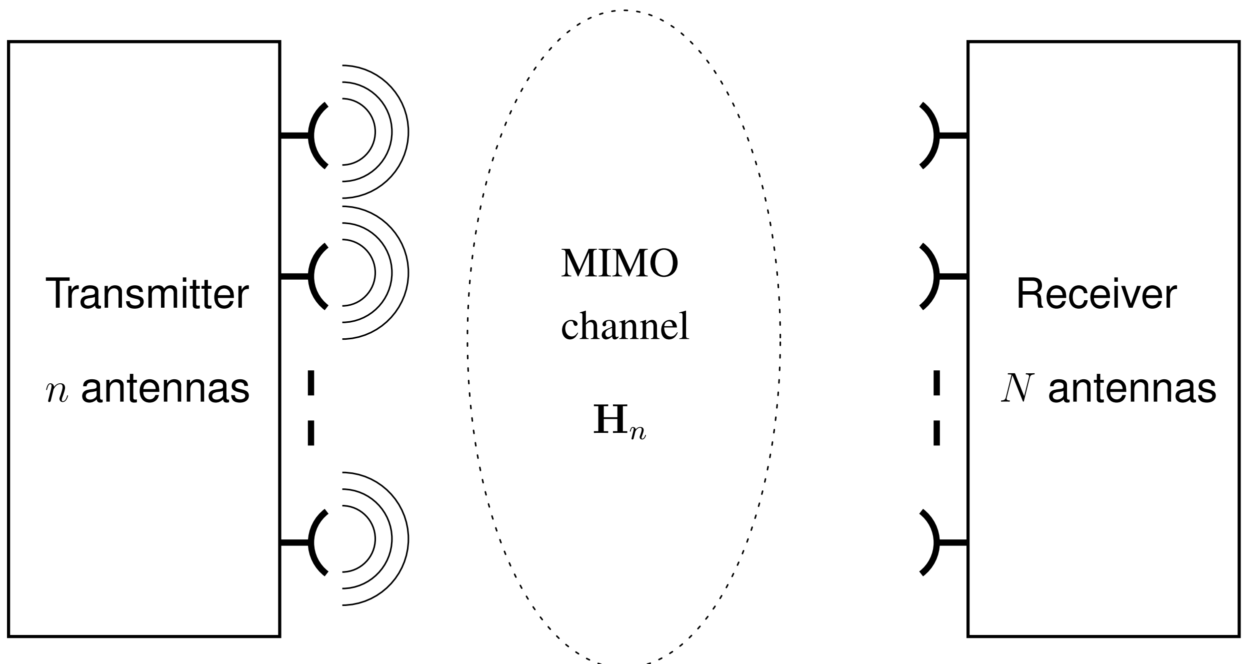

Figure 1: Multiple Input Multiple Output (MIMO) wireless communication

Shannon's mutual information

Shannon's mutual information per receive antenna of the $N \times n$ random MIMO channel $\mathbf{H}_n$:

$$C_n\left(\varsigma^2\right) = \frac{1}{N} \mathbb{E} \log\det \left(\mathbf{I}_N + \frac{1}{\varsigma^2} \mathbf{H}_n \mathbf{H}_n^{\mathrm{H}} \right)$$

where ς^2 is a known parameter (noise variance).

Information theory: $NC_n\left(\varsigma^2\right)$ is the maximum data rate attainable by the transmission system.

Behaviour of $C_n\left(\varsigma^2\right)$ as $n \to \infty$ and $\frac{N}{n} \to c > 0$?

Spectral measure and Stieltjes Transform

- $C_n\left(\varsigma^2\right) = \mathbb{E}\frac{1}{N}\sum_{i=1}^{N}\log\left(1+\frac{\lambda_{i,n}}{\varsigma^2}\right) = \mathbb{E}\int \log\left(1+\frac{t}{\varsigma^2}\right)\mu_n(dt)$ where μ_n is the spectral measure (empirical distribution of eigenvalues $\{\lambda_{1,n},\ldots,\lambda_{N,n}\}$) of $\mathbf{H}_n\mathbf{H}_n^{\mathrm{H}}$.
- Given a certain statistical model for $\mathbf{H}_n$, one hopes that the spectral measure μ_n converges weakly to a deterministic Limit Spectral probability Measure (LSM) μ, in order to have

$$C_n\left(\varsigma^2\right) \xrightarrow[n\to\infty]{} C^*\left(\varsigma^2\right) = \int \log\left(1+\frac{t}{\varsigma^2}\right)\mu(dt)\ .$$

- We study μ_n in the asymptotic regime, or equivalently, its Stieltjes Transform (ST)

$$\mathbf{f}_{\mu_n}(z) = \int \frac{1}{t-z}\mu_n(dt)\ .$$

- Weak convergence of μ_n towards μ is equivalent to convergence of $\mathbf{f}_{\mu_n}(z)$ towards the ST $\mathbf{f}_{\mu}(z)$ of the LSM μ.

Channel statistical model 1

"Ricean" Channel Model

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

- $\mathbf{Z}_n = \left[Z_{i,j}^{(n)} \right]$, elements of a Gaussian stationary two dimensional process with covariance function κ:

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

- $\mathbf{B}_n$ is a deterministic matrix (Rice component).

Channel statistical model 2

Channel matrix is $\mathbf{F}_N \mathbf{H}_n \mathbf{F}_n^{\mathrm{H}}$ where $\mathbf{F}_l$ is the $l \times l$ Fourier matrix and

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

- Elements of $\mathbf{Y}_n = \left[Y_{i,j}^{(n)} \right]$ written $Y_{i,j}^{(n)} = \dfrac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}$ with X_{ij} standard Gaussian independent random variables.
- $\mathbf{A}_n$ is a deterministic matrix.

Sometimes we shall assume:

(A) Variance profile is $\sigma_{ij}(n)^2 = \sigma^2 \left(\dfrac{i}{N}, \dfrac{j}{n} \right)$ where $\sigma^2(x, y)$ is a continuous function on $[0, 1]^2$ called a limit variance profile.

Link between models 1 and 2

For asymptotic study, model 1 can be replaced with model 2 with

- Assumption **(A)** with $\sigma^2(x,y) = \Gamma(x,y)$ where

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

  is the Spectral Density of the process $Z_{i,j}$.

- $\mathbf{A}_n$ is the two-dimensional Fourier Transform of $\mathbf{B}_n$.

and some assumptions.
Argument formalized in Hachem, Loubaton and Najim'05.

Problem statement

Model 2: $\mathbf{H}_n = \mathbf{Y}_n + \mathbf{A}_n$ with size $N \times n$.

- $\mathbf{Y}_n = \left[Y_{i,j}^{(n)} \right]$ with $Y_{i,j}^{(n)} = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}$, random variables X_{ij} are centered unit variance iid.
  We release Gaussianity assumption on X_{ij}.
- $\mathbf{A}_n$ is a deterministic matrix.

With appropriate additional assumptions,

- Characterize the asymptotic behaviour of the spectral measure μ_n of $\mathbf{H}_n \mathbf{H}_n^{\mathrm{H}}$ as $n \to \infty$ and $N/n \to c > 0$, or equivalently, its ST $\mathbf{f}_{\mu_n}(z)$.
- Deduce the asymptotic behaviour of Shannon's mutual information $C_n \left(\varsigma^2 \right)$.

Some particular cases

$\boxed{\textsc{The centered case } (\mathbf{A}_n = \mathbf{0})}$

Assume **(A)**, *i.e.*, $\exists$ a limit variance profile.

- Girko'90: μ_n converges weakly to a deterministic probability measure μ which ST $\mathbf{f}_\mu(z)$ has the form $\mathbf{f}_\mu(z) = \int_0^1 p(u,z)du$.
  Function $p(u,z)$ continuous in u for every z, ST of a probability measure in z for every u, defined as the unique solution of an implicit equation.

- Same result can be deduced from the work of Boutet de Monvel, Khorunzhyi and Vasilchuck (96).

- And also from Shlyakhtenko's (96) result stated for Wigner Gaussian matrices. His approach based on the concept of freeness with amalgamation.

Remark on the general non centered case

Even if we have a limit variance profile $\sigma^2(x, y)$ for the elements of $\mathbf{Y}_n$ and if $\mathbf{A}_n\mathbf{A}_n^{\mathrm{H}}$ has a limit spectral measure, the spectral measure μ_n of $\mathbf{H}_n\mathbf{H}_n^{\mathrm{H}}$ **does not converge** except in some very **specific cases**.

SPECIFIC CASE 1: $\sigma(x,y)$ CONSTANT AND $\mathbf{A}\mathbf{A}^{\mathrm{H}}$ HAS A LSM

Case

- $\sigma(x,y) = \sigma$ is a constant, *i.e.*, $\mathbf{Y}_n$ has iid elements,
- The spectral measure ν_n of $\mathbf{A}_n\mathbf{A}_n^{\mathrm{H}}$ converges weakly

$$\nu_n \Longrightarrow \nu$$

treated by Brent Dozier and Silverstein (04): μ_n converges to a deterministic probability measure which ST $\mathbf{f}(z)$ is the unique solution to

$$\mathbf{f}(z) = \int \frac{\nu(dt)}{-z\left(1 + c\sigma^2\mathbf{f}(z)\right) + (1-c)\sigma^2 + \frac{t}{1 + c\sigma^2\mathbf{f}(z)}}$$

in the class of ST of probability measures over $\mathbb{R}_+$.

SPECIFIC CASE 2: $\sigma^2(x,y)$ NON TRIVIAL AND $\mathbf{A}$ DIAGONAL

Hachem, Loubaton, Najim'05:

- Existence of a limit variance profile **(A)**.
- Moment assumption: $\exists\, \varepsilon > 0$ where $\mathbb{E}\,|X_{ij}|^{4+\varepsilon} < \infty$.
  Can be lightened by a truncation argument (Bai and Silverstein).
- $\mathbf{A}_n$ diagonal, *i.e.*, when $n \geq N$ (which we shall assume), has the form

$$\mathbf{A}_n = \begin{bmatrix} A_{11} & & \cdots & \cdots & 0 \\ & \ddots & & & \vdots \\ 0 & & A_{NN} & \cdots & 0 \end{bmatrix}$$

- $\displaystyle \frac{1}{N}\sum_{i=1}^{N} \delta_{\left(i/N, |A_{ii}|^2\right)} \Longrightarrow H(dt, d\lambda)$, compactly supported pr. measure in $[0,1] \times \mathbb{R}_+$.

  "Stonger" than convergence of the empirical distributon $\displaystyle \frac{1}{N}\sum_{i=1}^{N} \delta_{|A_{ii}|^2}$.

Specific case 2: technique

- Resolvent is $\mathbf{Q}_n(z) = \left(\mathbf{H}_n\mathbf{H}_n^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}$. ST associated with the spectral measure μ_n of $\mathbf{H}_n\mathbf{H}_n^{\mathrm{H}}$:

$$\mathbf{f}_{\mu_n}(z) = \int \frac{1}{t-z}\mu_n(dt) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_{i,n}-z} = \frac{1}{N}\mathrm{tr}\mathbf{Q}_n(z)$$

- Let $\tilde{\mu}_N$ be the spectral measure of $\mathbf{H}_n^{\mathrm{H}}\mathbf{H}_n$. Associated ST is $\mathbf{f}_{\tilde{\mu}_n}(z) = \frac{1}{N}\mathrm{tr}\widetilde{\mathbf{Q}}_n(z)$ with $\widetilde{\mathbf{Q}}_n(z) = \left(\mathbf{H}_n^{\mathrm{H}}\mathbf{H}_n - z\mathbf{I}_n\right)^{-1}$.

- We study jointly the convergence of $\mathbf{f}_{\mu_n}$ and $\mathbf{f}_{\tilde{\mu}_n}$ by considering the diagonal terms $Q_{ii}(z)$ and $\widetilde{Q}_{jj}(z)$ of $\mathbf{Q}_n(z)$ and $\widetilde{\mathbf{Q}}_n(z)$.

Specific case 2: technique

- We establish convergence of measures

$$
\begin{aligned}
L_n(z, du, d\lambda) &= \frac{1}{N}\sum_{i=1}^{N} Q_{ii}(z)\ \delta_{\left(\frac{i}{N}, |A_{ii}|^2\right)}(du, d\lambda) \\
\widetilde{L}_n(z, du, d\lambda) &= \frac{1}{n}\sum_{j=1}^{N} \widetilde{Q}_{jj}(z)\ \delta_{\left(\frac{j}{n}, |A_{jj}|^2\right)}(du, d\lambda) \\
&\qquad + \frac{1}{n}\sum_{j=N+1}^{n} \widetilde{Q}_{jj}(z)\ \delta_{\frac{j}{n}}(du) \otimes \delta_0(d\lambda)
\end{aligned}
$$

SPECIFIC CASE 2: LIMIT SPECTRAL MEASURE

- Consider the following system: for every bounded continuous g,

$$\int g\,d\pi(z,du,d\lambda) = \int \frac{g(u,\lambda)}{-z - z\int \sigma^2(u,t)d\tilde{\pi}(z,dt,d\zeta) + \frac{\lambda}{1+c\int \sigma^2(t,cu)d\pi(z,dt,d\zeta)}} H(du,d\lambda)$$

$$\int g\,d\tilde{\pi}(z,du,d\lambda) = c\int \frac{g(cu,\lambda)}{-z - cz\int \sigma^2(t,cu)d\pi(z,dt,d\zeta) + \frac{\lambda}{1+\int \sigma^2(u,t)d\tilde{\pi}(z,dt,d\zeta)}} H(du,d\lambda)$$

$$+ (1-c)\int_c^1 \frac{g(u,0)}{-z - cz\int \sigma^2(t,u)d\pi(z,dt,d\zeta)}\,du$$

System has a unique solution $(\pi,\tilde{\pi})$ in a certain class of complex measures (the Stieltjes kernels).

- π and $\tilde{\pi}$ are the limits of L_n and $\widetilde{L}_n$ in the weak convergence of complex measures.
- The limit ST $\mathbf{f}_\mu$ and $\mathbf{f}_{\tilde{\mu}}$ are then

$$\mathbf{f}_\mu(z) = \int \pi(z,dt,d\lambda) \quad \text{and} \quad \mathbf{f}_{\tilde{\mu}}(z) = \int \tilde{\pi}(z,dt,d\lambda)$$

The general case

- We assume $\sigma^2(x,y)$ non trivial and $\mathbf{A}_n$ has no particular structure.
- Difficult to devise simple conditions for the existence of a limit spectral measure, *i.e.*, an "extension" of assumption $\frac{1}{N}\sum_{i=1}^{N}\delta_{(i/N,|A_{ii}|^2)} \Longrightarrow H(dt,d\lambda)$ that we used for the case $\mathbf{A}_n$ is diagonal.
- An alternative approach: look for a deterministic approximation of the empirical ST: there exists a a $N \times N$ deterministic matrix function $\mathbf{T}_n(z)$ such that

$$\mathbf{f}_{\mu_n}(z) - \frac{1}{N}\mathrm{tr}\mathbf{T}_n(z) \xrightarrow[n\to\infty]{} 0 \text{ almost surely}$$

  This "deterministic approximation" dates back to Girko.

Deterministic approximation: assumptions

Hachem, Loubaton, Najim'05 (preprint): Extension of Girko's result and simplification of his proof, approximation of Shannon's mutual information.

Problem: approximate the spectral measure of $\mathbf{H}_n = \mathbf{Y}_n + \mathbf{A}_n$ with

- $Y_{i,j}^{(n)} = \frac{\sigma_{ij}(n)}{\sqrt{n}} X_{ij}$ with X_{ij} centered unit variance iid and $\mathbb{E}\,|X_{11}|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$. Last assumption can be lightened.
- $\sup_{i,j,n} \sigma_{ij}^2(n) < \infty$.
- Euclidean norms of rows and columns of $\mathbf{A}_n$ uniformly bounded.

Girko assumed boundedness of ℓ_1 norms of rows and columns.
In wireless communications, columns of $\mathbf{A}_n$ have typically the form

$$\frac{C}{\sqrt{N}}\left[1, \exp(\imath\omega), \ldots, \exp(\imath(N-1)\omega)\right]^{\mathrm{T}}$$

ℓ_1 norm increases in $\sqrt{N}$ while Euclidean (ℓ_2) norm is bounded.

Deterministic approximation: result

Let $\mathbf{D}^{(j)} = \mathrm{diag}\left(\left[\sigma_{1j}^2, \ldots, \sigma_{Nj}^2\right]\right)$ and $\widetilde{\mathbf{D}}^{(i)} = \mathrm{diag}\left(\left[\sigma_{i1}^2, \ldots, \sigma_{in}^2\right]\right)$.

- The deterministic system of $N+n$ equations:

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

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

where

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

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

admits a unique solution $(\psi^{(1)}, \ldots, \psi^{(N)}, \tilde{\psi}^{(1)}, \ldots, \tilde{\psi}^{(n)})$ in the class of Stieltjes Transforms of probability measures over $\mathbb{R}_+$.

Deterministic approximation: result

- Almost surely,

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

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

Back to mutual information

Mutual information can be written

$$C_n\left(\varsigma^2\right) = \int_{\varsigma^2}^{\infty}\left(\frac{1}{\omega} - \mathbb{E}\frac{1}{N}\mathrm{tr}\mathbf{Q}_n\left(-\omega\right)\right)d\omega$$

Combining this expression with the last result, we can establish:
Let

$$\begin{aligned}\overline{C}_n(\varsigma^2) = \frac{1}{N}\log\det\left[\frac{\mathbf{\Psi}\left(-\varsigma^2\right)^{-1}}{\varsigma^2} + \mathbf{A}\widetilde{\mathbf{\Psi}}\left(-\varsigma^2\right)\mathbf{A}^{\mathrm{H}}\right] \\ + \frac{1}{N}\log\det\frac{\widetilde{\mathbf{\Psi}}\left(-\varsigma^2\right)^{-1}}{\varsigma^2} - \frac{\varsigma^2}{Nn}\sum_{\substack{i=1:N\\ j=1:n}}\sigma_{ij}^2 T_{ii}(-\varsigma^2)\widetilde{T}_{jj}(-\varsigma^2)\end{aligned}$$

where T_{ii} and $\widetilde{T}_{jj}$ are the diagonal elements of $\mathbf{T}_n(z)$ and $\widetilde{\mathbf{T}}_n(z)$. Then

$$C_n\left(\varsigma^2\right) - \overline{C}_n(\varsigma^2) \xrightarrow[n\to\infty]{} 0.$$

General case: main steps of the proof

$\boxed{\text{STEP 1: EXISTENCE AND UNICITY OF } \mathbf{T}(z)}$

Existence and unicity of $\mathbf{T}_n(z)$ and $\widetilde{\mathbf{T}}_n(z)$ as solutions of the system of $N + n$ equations above.

- Existence by an iterative scheme.
- Unicity in a certain region of $\mathbb{C}$. In $\mathbb{C} - \mathbb{R}_+$ by analytic continuation.
- Use an extension of complex analysis results about Stieltjes transforms of probability measures over $\mathbb{R}_+$: let $\mathbf{T}(z)$ be an analytical matrix function on $\mathbb{C}_+ = \{z : \Im z > 0\}$ such that $\Im \mathbf{T}(z) \geq 0$ on $\mathbb{C}_+$ and $\Im z\mathbf{T}(z) \geq 0$ on $\mathbb{C}_+$, (as non negative matrices). Then there exists a matrix $\mathbf{C} \geq \mathbf{0}$ and a matrix valued measure $\boldsymbol{\mu}$ carried by $\mathbb{R}_+$ such as $\boldsymbol{\mu}(A) \geq 0$ for every Borel set A of $\mathbb{R}_+$, and

$$\mathbf{T}(z) = \mathbf{C} + \int \frac{1}{t - z} \boldsymbol{\mu}(dt) \quad \text{with} \quad \mathrm{tr} \int \frac{1}{1 + t} \boldsymbol{\mu}(dt) < \infty$$

STEP 2: INTRODUCING NEW FUNCTIONS $\mathbf{R}(z)$ AND $\widetilde{\mathbf{R}}(z)$

We introduce intermediate matrices $\mathbf{R}_n(z)$ and $\widetilde{\mathbf{R}}_n(z)$ defined as:

$$b^{(i)}(z) = \frac{-1}{z\left(1 + \frac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{D}}^{(i)}\widetilde{\mathbf{Q}}(z)\right)\right)}, \qquad \mathbf{B}(z) = \mathrm{diag}\left(\left[b^{(1)}(z), \ldots, b^{(N)}(z)\right]\right),$$

$$\tilde{b}^{(j)}(z) = \frac{-1}{z\left(1 + \frac{1}{n}\mathrm{tr}\left(\mathbf{D}^{(j)}\mathbf{Q}(z)\right)\right)}, \qquad \widetilde{\mathbf{B}}(z) = \mathrm{diag}\left(\left[\tilde{b}^{(1)}(z), \ldots, \tilde{b}^{(n)}(z)\right]\right),$$

$$\mathbf{R}(z) = \left(\mathbf{B}^{-1}(z) - z\mathbf{A}\widetilde{\mathbf{B}}(z)\mathbf{A}^{\mathrm{H}}\right)^{-1}, \quad \widetilde{\mathbf{R}}(z) = \left(\widetilde{\mathbf{B}}^{-1}(z) - z\mathbf{A}^{\mathrm{H}}\mathbf{B}(z)\mathbf{A}\right)^{-1}.$$

STEP 2: INTRODUCING NEW FUNCTIONS $\mathbf{R}(z)$ AND $\widetilde{\mathbf{R}}(z)$

We show that for any diagonal matrices $\mathbf{U}_n$ and $\widetilde{\mathbf{U}}_n$ such that $\sup_n \|\mathbf{U}_n\| < \infty$ and $\sup_n \|\widetilde{\mathbf{U}}_n\| < \infty$, we have on $\mathbb{C}_+$,

$$\begin{array}{lcl} \mathbb{E}\left|\frac{1}{n}\mathrm{tr}\left(\left(\mathbf{Q}_n(z) - \mathbf{R}_n(z)\right)\mathbf{U}_n\right)\right|^{2+\varepsilon/2} & < & \mathsf{Cst} \times n^{-(1+\varepsilon/4)} \quad \text{and} \\ \mathbb{E}\left|\frac{1}{n}\mathrm{tr}\left(\left(\widetilde{\mathbf{Q}}_n(z) - \widetilde{\mathbf{R}}_n(z)\right)\widetilde{\mathbf{U}}_n\right)\right|^{2+\varepsilon/2} & < & \mathsf{Cst} \times n^{-(1+\varepsilon/4)} \end{array}$$

Derivations along the lines of those of Brent Dozier and Silverstein (04).
Bai and Silverstein's (98) lemma is of prime importance: in our context, for any $p \geq 2$,

$$\mathbb{E}\left|\frac{1}{N}\mathbf{x}_N^{\mathrm{H}}\mathbf{Z}_N\mathbf{x}_N - \frac{1}{N}\mathrm{tr}\mathbf{Z}_N\right|^p < \frac{\mathsf{Cst}}{N^{p/2}}$$

for $\mathbf{x}_N = [X_1, \ldots, X_N]^{\mathrm{T}}$ with X_i iid centered unit variance random variables with $\mathbb{E}|X_{11}|^{2p} < \infty$, and $\mathbf{Z}_N$ is a $N \times N$ random matrix independent of $\mathbf{x}_N$ such that $\sup_N \|\mathbf{Z}_N\| < \infty$.

Step 3: $\frac{1}{n}\mathrm{tr}\mathbf{R}$ is close to $\frac{1}{n}\mathrm{tr}\mathbf{T}$

We show that in a certain region $\mathcal{D}$ of $\mathbb{C}_+$,

$$\mathbb{E}\left|\frac{1}{n}\mathrm{tr}\left(\mathbf{R}(z)-\mathbf{T}(z)\right)\right|^{2+\varepsilon/2} < \frac{\mathsf{Cst}}{n^{1+\varepsilon/4}} \quad \text{and}$$

$$\mathbb{E}\left|\frac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{R}}(z)-\widetilde{\mathbf{T}}(z)\right)\right|^{2+\varepsilon/2} < \frac{\mathsf{Cst}}{n^{1+\varepsilon/4}}$$

Idea:

Recall that $b^{(i)}(z) = \dfrac{-1}{z\left(1+\frac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{D}}^{(i)}\widetilde{\mathbf{Q}}(z)\right)\right)}$.

From step 2 with $\widetilde{\mathbf{U}} = \widetilde{\mathbf{D}}^{(i)}$ we have $\dfrac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{D}}^{(i)}\widetilde{\mathbf{Q}}(z)\right) = \dfrac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{D}}^{(i)}\widetilde{\mathbf{R}}(z)\right) + \epsilon^{(i)}$.

It results that $b^{(i)}(z) = \dfrac{-1}{z\left(1+\frac{1}{n}\mathrm{tr}\left(\widetilde{\mathbf{D}}^{(i)}\widetilde{\mathbf{R}}(z)\right)\right)} + \underline{\epsilon}^{(i)}$ with

$\mathbb{E}\left|\underline{\epsilon}^{(i)}\right|^{2+\varepsilon/2} < \mathsf{Cst} \times n^{-(1+\varepsilon/4)}$.

STEP 3: $\frac{1}{n}\mathrm{tr}\mathbf{R}$ IS CLOSE TO $\frac{1}{n}\mathrm{tr}\mathbf{T}$

Similarly $\tilde{b}^{(j)}(z) = \dfrac{-1}{z\left(1+\frac{1}{n}\mathrm{tr}\left(\mathbf{D}^{(j)}\mathbf{R}(z)\right)\right)} + \underline{\tilde{\epsilon}}^{(j)}$.

Recall that

$$\mathbf{R}(z) = \left(\mathbf{B}^{-1}(z) - z\mathbf{A}\widetilde{\mathbf{B}}(z)\mathbf{A}^{\mathrm{H}}\right)^{-1} \quad \text{and} \quad \widetilde{\mathbf{R}}(z) = \left(\widetilde{\mathbf{B}}^{-1}(z) - z\mathbf{A}^{\mathrm{H}}\mathbf{B}(z)\mathbf{A}\right)^{-1}.$$

So, up to the $\underline{\epsilon}^{(i)}$ and $\underline{\tilde{\epsilon}}^{(j)}$, matrices $\mathbf{B}$ and $\widetilde{\mathbf{B}}$ satisfy the same system as $\mathbf{\Psi}$ and $\widetilde{\mathbf{\Psi}}$.

With this idea, $\left(\mathbf{R}, \widetilde{\mathbf{R}}\right)$ can be approached by $\left(\mathbf{T}, \widetilde{\mathbf{T}}\right)$ for z carefully chosen (in the region $\mathcal{D}$).

Putting pieces together

Step 1: $\mathbf{T}_n(z)$ and $\widetilde{\mathbf{T}}_n(z)$ exist and are unique as solutions of a system of equations.

Step 2: $\mathbb{E}\left|\frac{1}{n}\mathrm{tr}\left((\mathbf{Q}_n(z) - \mathbf{R}_n(z))\right)\right|^{2+\varepsilon/2} < \mathsf{Cst} \times n^{-(1+\varepsilon/4)}$ by taking $\mathbf{U}_n = \mathbf{I}_N$.

Step 3: $\mathbb{E}\left|\frac{1}{n}\mathrm{tr}\left(\mathbf{R}_n(z) - \mathbf{T}_n(z)\right)\right|^{2+\varepsilon/2} < \mathsf{Cst} \times n^{-(1+\varepsilon/4)}$ in a region $\mathcal{D}$.

Consequence: $\frac{1}{N}\mathrm{tr}\left(\mathbf{Q}_n(z) - \mathbf{T}_n(z)\right) \xrightarrow[n\to\infty]{} 0$ almost surely on $\mathbb{C} - \mathbb{R}_+$ by Borel-Cantelli's lemma and by analytic continuation.

Towards a Central Limit Theorem

Let $I_n\left(\varsigma^2\right) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{\varsigma^2}\mathbf{H}_n\mathbf{H}_n^{\mathrm{H}}\right)$ so that $C_n\left(\varsigma^2\right) = \mathbb{E} I_n\left(\varsigma^2\right)$.

- CLT over I_n as $n \to \infty$ and $N/n \to c > 0$, at least in some particular cases such as $\mathbf{A}_n = \mathbf{0}$ in the model $\mathbf{H}_n = \mathbf{Y}_n + \mathbf{A}_n$. We shall assume this case.
- By means of the "Gaussian approximation", we have an idea of the "outage probability" $\mathbb{P}\left(I_n < \text{a given threshold } R\right)$.
  In certain situations, this gives the probability that the channel cannot provide data rate R.
- Two terms :
  - CLT over $\chi_{1,n} = N\left(I_n - C_n\right)$ and variance derivation.
  - Bias $\chi_{2,n} = N\left(C_n - \overline{C}_n\right)$ between mutual information NC_n and the deterministic approximation $N\overline{C}_n$.

THE TERM $\chi_{1,n}$

Approach: CLT for martingales as in Girko and in Bai and Silverstein'04.

Notations:

- $\mathbf{Y}^{(j)}$ is the $N \times (n-1)$ matrix that remains after extracting column j denoted as $\mathbf{y}^{(j)}$ from $\mathbf{Y}$.
- $\mathbf{Q}^{(j)}(z)$ is the resolvent $\mathbf{Q}^{(j)}(z) = \left(\mathbf{Y}^{(j)} {\mathbf{Y}^{(j)}}^{\mathrm{H}} - z\mathbf{I}_n\right)^{-1}$.
- $\mathcal{F}^{(j)}$ is the $\sigma-$field $\mathcal{F}^{(j)} = \sigma\left(\mathbf{y}^{(j)}, \ldots, \mathbf{y}^{(n)}\right)$.
- $\mathbb{E}^{(j)}$ is the conditional expectation $\mathbb{E}\left[\,.\,\|\mathcal{F}^{(j)}\right]$.
- $I_n^{(j)}\left(\varsigma^2\right) = \dfrac{1}{N}\log\det\left(\mathbf{I}_N + \dfrac{1}{\varsigma^2}\mathbf{Y}_n^{(j)} {\mathbf{Y}_n^{(j)}}^{\mathrm{H}}\right)$.

The term $\chi_{1,n}$

- We have

$$N\left(I_n - \mathbb{E}I_n\right) = N\sum_{j=1}^{n}\left(\mathbb{E}^{(j)} - \mathbb{E}^{(j+1)}\right) I_n$$
$$= N\sum_{j=1}^{n}\left(\mathbb{E}^{(j)} - \mathbb{E}^{(j+1)}\right)\left(I_n - I_n^{(j)}\right) \quad \text{due to } \mathbb{E}^{(j)} I_n^{(j)} = \mathbb{E}^{(j+1)} I_n^{(j)}.$$

- By standard matrix manipulations, we have

$$N\left(I_n - I_n^{(j)}\right) = \log\left(\varsigma^2\right) + \log\left(1 + \mathbf{y}^{(j)\mathrm{H}}\mathbf{Q}^{(j)}\left(-\varsigma^2\right)\mathbf{y}^{(j)}\right)$$

- Sequence $\gamma^{(j)} = \left(\mathbb{E}^{(j)} - \mathbb{E}^{(j+1)}\right)\log\left(1 + \mathbf{y}^{(j)\mathrm{H}}\mathbf{Q}^{(j)}\left(-\varsigma^2\right)\mathbf{y}^{(j)}\right)$ is a martingale difference sequence with respect to the increasing filtration $\mathcal{F}^{(n)}, \ldots, \mathcal{F}^{(1)}$. Apply the CLT for martingales to $\sum_{j=1}^{n}\gamma^{(j)}$.

- Variance of $\chi_{1,n}$ is $\mathcal{O}(1)$.

THE BIAS TERM $\chi_{2,n}$

$$\chi_{2,n} = N\left(C_n - \overline{C}_n\right)$$

We get back to ST by taking the derivative with respect to ς^2:

$$\frac{d\chi_{2,n}}{d\varsigma^2} = -\mathrm{tr}\left(\mathbb{E}\mathbf{Q}_n\left(-\varsigma^2\right) - \mathbf{T}_n\left(-\varsigma^2\right)\right)$$

We obtain

$$\frac{d\chi_{2,n}}{d\varsigma^2} \xrightarrow[n\to\infty]{} \left(\mathbb{E}\left|X_{11}\right|^4 - 2\right) \times \text{Cst}$$

$\chi_{2,n} \to 0$ in the case elements of $\mathbf{Y}_n$ are Gaussian.