Random matrices and communication systems: WEEK 4

2bis H is random and varying ergodically over time (fast fading)

Let us consider again the expression found for the ergodic capacity, under the assumption that the fading coefficients h_{jk} are i.i.d.$\sim \mathcal{N}_{\mathbb{C}}(0,1)$ random variables:

$$C_{\text{erg}} = \mathbb{E}\left(\log\det\left(I + \frac{P}{n}\,HH^*\right)\right)$$

This expression can be computed explicitly via random matrix theory, which is the subject of this course. Before that, we will see below a relatively simple computation that will provide us with a lower bound on C_{erg}, that turns out to be asymptotically tight, either when $n \to \infty$ or $P \to \infty$.

As a preliminary, we need the *Paley-Zygmund inequality*: if $X \geq 0$ is a square-integrable random variable, then

$$\mathbb{P}(X > t) \geq \frac{(\mathbb{E}(X) - t)^2}{\mathbb{E}(X^2)} \quad \forall 0 \leq t \leq \mathbb{E}(X)$$

The proof of this fact is an application of Cauchy-Schwarz' inequality.

Let now $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of the (positive semi-definite) $n \times n$ matrix $\frac{1}{n}\,HH^*$. This allows us to write

$$C_{\text{erg}} = \mathbb{E}\left(\sum_{j=1}^n \log(1 + P\lambda_j)\right)$$

Let furthermore λ be one of the eigenvalues $\lambda_1, \ldots, \lambda_n$ picked uniformly at random. We obtain

$$C_{\text{erg}} = n\,\mathbb{E}(\log(1+P\lambda)) \geq n\,\log(1+Pt)\,\mathbb{P}(\lambda > t) \geq n\,\log(1+Pt)\,\frac{(\mathbb{E}(\lambda)-t)^2}{\mathbb{E}(\lambda^2)}$$

for all $0 \leq t \leq \mathbb{E}(\lambda)$, by the above Paley-Zygmund inequality. Let us compute (these are by the way our first random matrix computations):

$$\mathbb{E}(\lambda) = \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n \lambda_j\right) = \mathbb{E}\left(\frac{1}{n}\mathrm{Tr}\left(\frac{1}{n}\,HH^*\right)\right) = \frac{1}{n^2}\sum_{j,k=1}^n \mathbb{E}\left(|h_{jk}|^2\right) = \frac{1}{n^2}\,n^2 = 1$$

and

$$\mathbb{E}\left(\lambda^2\right) = \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n \lambda_j^2\right) = \mathbb{E}\left(\frac{1}{n}\mathrm{Tr}\left(\left(\frac{1}{n}\,HH^*\right)^2\right)\right) = \frac{1}{n^3}\sum_{j,k,l,m=1}^n \mathbb{E}\left(h_{jk}\,\overline{h_{lk}}\,h_{lm}\overline{h_{jm}}\right)$$

Notice that because the h_{jk} are i.i.d. and $\mathbb{E}(h_{jk}) = 0$ for all j,k, it holds that $\mathbb{E}\left(h_{jk}\,\overline{h_{lk}}\,h_{lm}\overline{h_{jm}}\right) = 0$ unless $j = l$ or $k = m$. We therefore obtain

$$\begin{aligned}
\mathbb{E}\left(\lambda^2\right) &= \frac{1}{n^3}\left(\sum_{j,k=1}^n \mathbb{E}\left(|h_{jk}|^4\right) + \sum_{j\; k\neq m} \mathbb{E}\left(|h_{jk}|^2\,|h_{jm}|^2\right) + \sum_{j\neq l,\, k} \mathbb{E}\left(|h_{jk}|^2\,|h_{lk}|^2\right)\right) \\
&= \frac{1}{n^3}\left(2n^2 + n^2\,(n-1) + n^2\,(n-1)\right) = 2
\end{aligned}$$

This finally implies that for all $0 \leq t \leq 1$,

$$C_{\text{erg}} \geq n\,\log(1+Pt)\,\frac{(1-t)^2}{2} = \frac{n}{8}\,\log(1+P/2)$$

by choosing $t = 1/2$. This is to be compared with the case of a fixed deterministic matrix H_0 whose coefficients are all equal to 1, with corresponding capacity $C_0 = \log(1 + P\,n^2)$ (see last lecture). It holds that $\mathbb{E}(|h_{jk}|^2) = 1 = (H_0)_{jk}$ for all j,k, but notice that:

a) for fixed P and $n \to \infty$, $C_0 \simeq \log n$, while $C_{\text{erg}} \gtrsim n$.

b) for fixed n and $P \to \infty$, $C_0 \simeq 1 \log P$, while $C_{\text{erg}} \gtrsim n \log P$

So in multiple antenna systems, random (i.i.d.) fading actually improves the capacity, contrary to single antenna systems.

3 H is random and fixed over time (slow fading)

We assume now that the matrix H admits the continuous pdf $p_H(\cdot)$, whose support contains the all zero matrix, that H is fixed over time, and that its realization is known at the receiver, but not at the transmitter. In this case, the capacity of the channel is zero and the single-letter characterization of its outage probability reads:

$$P_{\text{out}}(R) = \inf_{p_X \,:\, \text{Tr}(Q_X) \leq P} \mathbb{P}_H(I(X;Y) < R)$$

where R is the target rate chosen by the transmitter. As we have already seen, for a any given H,

$$I(X;Y) \leq \log\det(I + H\,Q_X\,H^*)$$

and the equality is met when $X \sim \mathcal{N}_{\mathbb{C}}(0, Q_X)$ (that does not depend on the particular realization of H). So

$$P_{\text{out}}(R) = \inf_{Q_X \geq 0 \,:\, \text{Tr}(Q_X) \leq P} \mathbb{P}_H(\log\det(I + H\,Q_X\,H^*) < R)$$

Let us now consider the particular case where the matrix H has i.i.d entries $h_{jk} \sim \mathcal{N}_{\mathbb{C}}(0,1)$. In this case, H and HU share the same distribution, for any deterministic $n \times n$ unitary matrix U, so we may as well take Q_X diagonal in the above optimization problem. Therefore,

$$P_{\text{out}}(R) = \inf_{\substack{d_1,\ldots,d_n \geq 0 \\ \sum_{j=1}^n d_j \leq P}} \mathbb{P}_H(\log\det(I + HDH^*) < R)$$

where $D = \text{diag}(d_1, \ldots, d_n)$. Solving further this optimization problem turns out to be difficult. Notice first that the answer depends on the target rate R:

- if R is (sufficiently) small, then setting $D = \frac{P}{n} I$ achieves the minimum outage probability, as in this case, the law of large numbers plays for us: $\log\det(I + \frac{P}{n} HH^*)$ is highly likely to be around its average value and therefore to exceed R.

- if R is (sufficiently) large, then the law of large numbers plays on the contrary against us, and it is therefore better in this case to put "all our eggs in one basket", that is, to use $D = \text{diag}(P, 0, \ldots, 0)$ and hope that the chosen channel is by chance a good one that will prevent the mutual information to fall below the target rate.

Besides, it has been conjectured that in general, the optimal matrix D should be of the form

$$D = \text{diag}\Bigg(\underbrace{\frac{P}{k}, \ldots, \frac{P}{k}}_{k \text{ times}}, 0, \ldots, 0\Bigg)$$

where $1 \leq k \leq n$ is some integer parameter. We will see in the course how to analyze further this outage probability, with the help of random matrix theory.