# Random matrix theory : lecture 7

## Capacity of multi-antenna channels (cont'd)

Reminder: we study the multi-antenna channel $X \to Y = HX + Z$

main assumptions : • H $m \times n$ random matrix

• H varies ergodically over time ("second scenario")

• at each time instant, H has the same distribution $p(H)$

• the realizations of H are known to the receiver, but not to the transmitter

Under these assumptions, we get the following capacity expression:

$$C = \max_{Q \geq 0 : \mathrm{Tr}\, Q \leq P} \mathbb{E}\left( \underbrace{\log \det (I + HQH^*)}_{:= \psi_H(Q)} \right)$$

(NB: in order to lighten the notation, we skip the H in $\mathbb{E}_H$ and also skip the X in $Q_X$)

We are going to see how to simplify the above maximization problem when the distribution of H is invariant under some transformations.

ref: Abbé - Telatar - Zheng, Allerton 2005

## Preliminary proposition:

× The map $A \mapsto \log\det A$ is concave on the set of positive-definite matrices $(A>0)$

Therefore, $Q \mapsto \varphi_H(Q) := \log\det(I + HQH^*)$

× is also concave on the set of non-negative definite matrices (if $Q \geq 0$, then $I + HQH^* > 0$).

## Lemma 1

If $H$ has the same distribution as $H\Sigma$ (notation: $H \sim H\Sigma$) for any $n\times n$ matrix $\Sigma$ of the form $\Sigma = \begin{pmatrix} \pm1 & & 0 \\ & \pm1 & \\ 0 & & \pm1 \ddots \end{pmatrix}$, then

× $$C = \max_{Q \text{ diag} \geq 0 :\, \operatorname{Tr} Q \leq P} \mathbb{E}(\varphi_H(Q))$$

## Proof

- By assumption, $\varphi_H(Q)$ has the same distribution as

$$\begin{aligned} \varphi_{H\Sigma}(Q) &= \log\det(I + (H\Sigma)Q(H\Sigma)^*) \\ &= \log\det(I + H(\Sigma Q \Sigma^*)H^*) \\ &= \varphi_H(\Sigma Q \Sigma^*) \end{aligned}$$

× So $\mathbb{E}(\varphi_H(Q)) = \mathbb{E}(\varphi_H(\Sigma Q \Sigma^*))$

$$= \frac{1}{2}\left(\mathbb{E}(\varphi_H(Q)) + \mathbb{E}(\varphi_H(\Sigma Q \Sigma^*))\right)$$

i.e. $\mathbb{E}(\psi_H(Q)) = \mathbb{E}\left(\frac{1}{2}(\psi_H(Q) + \psi_H(\Sigma Q \Sigma^*))\right)$

- Since $Q \longmapsto \psi_H(Q)$ is concave, we further obtain

$$\mathbb{E}(\psi_H(Q)) \leq \mathbb{E}\left(\psi_H\left(\tfrac{1}{2}(Q + \Sigma Q \Sigma^*)\right)\right) \quad \left[\begin{smallmatrix}\text{Jensen}\\ \text{inequality}\end{smallmatrix}\right]$$

for any matrix $\Sigma = \mathrm{diag}(\pm 1)$ [<u>NB</u>: $\Sigma^* = \Sigma$]

- Consider e.g. $\Sigma = \begin{pmatrix} -1 & & 0 \\ & +1 & \\ 0 & & \ddots \; +1 \end{pmatrix}$ :

$$\Sigma Q \Sigma^* = \begin{pmatrix} q_{11} & -q_{12} \cdots -q_{1n} \\ -q_{21} & \\ \vdots & Q_1 \\ -q_{n1} & \end{pmatrix}$$

so $\frac{1}{2}(Q + \Sigma Q \Sigma^*) = \begin{pmatrix} q_{11} & 0 \cdots 0 \\ 0 & \\ \vdots & Q_1 \\ 0 & \end{pmatrix} := \tilde{Q}$

- That is, for any given input covariance matrix $Q$, there exists another covariance matrix $\tilde{Q}$ (with same trace) such that $\mathbb{E}(\psi(Q)) \leq \mathbb{E}(\psi(\tilde{Q}))$ and $\tilde{Q}$ has off-diagonal elements of the first row & first column which are all equal to zero.

- Proceeding recursively, we can "erase" all the off-diagonal elements and therefore show that choosing $Q$ diagonal is sufficient to maximize the expectation. #

## Lemma 2

× If $H \sim H\Pi$ for any $n \times n$ permutation matrix $\Pi$,

then $C = \max_{\rho \in [-\frac{1}{n-1}, 1]} \mathbb{E}(\psi_H(Q_\rho))$

where $Q_\rho = \frac{P}{n}\begin{pmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{pmatrix}$ equal diagonal elements & equal off-diagonal elements

× [NB: $\mathrm{Tr}\, Q_\rho = \frac{P}{n} \cdot n = P$ and $Q_\rho \geq 0$ iff $\rho \in [-\frac{1}{n}, 1]$]

Proof (same idea as before)

- By assumption, $\psi_H(Q) \sim \psi_{H\Pi}(Q) = \psi_H(\Pi Q \Pi^*)$,

  so $\mathbb{E}(\psi_H(Q)) = \mathbb{E}\left(\frac{1}{2}(\psi_H(Q) + \psi_H(\Pi Q \Pi^*))\right)$

  $\leq \mathbb{E}\left(\psi_H\left(\frac{1}{2}(Q + \Pi Q \Pi^*)\right)\right)$ concavity & Jensen

- Similarly, let $\tilde{Q} := \frac{1}{n!} \sum_{\Pi \in P(n)} \Pi Q \Pi^*$

  ∟ set of permutation matrices [NB: $\Pi^* \neq \Pi$ in general!]

  Then $\mathbb{E}(\psi_H(Q)) \leq \mathbb{E}(\psi_H(\tilde{Q}))$

- But note that $\tilde{Q}$ has equal diagonal elements & equal off-diagonal elements. Ex. in the case $n=2$:

× $$\frac{1}{2}(Q + \underbrace{\Pi}_{=\begin{pmatrix}0&1\\1&0\end{pmatrix}} Q \Pi^*) = \frac{1}{2}\left(\begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22}\end{pmatrix} + \begin{pmatrix} q_{22} & q_{21} \\ q_{12} & q_{11}\end{pmatrix}\right)$$

$$= \begin{pmatrix} \frac{1}{2}(q_{11}+q_{22}) & \frac{1}{2}(q_{12}+q_{21}) \\ \frac{1}{2}(q_{12}+q_{21}) & \frac{1}{2}(q_{11}+q_{22}) \end{pmatrix}$$

- Because of the trace constraint and the constraint that $\tilde{Q}$ is non-negative definite, $\tilde{Q}$ is necessarily of the form given in the lemma.

#

Remarks:

- The above argument had been already used in [Telatar 95], in the case where the optimal $Q$ was known to be diagonal.
- The same argument could have been used in lemma 1 also, observing that $\tilde{Q} := \frac{1}{2^n} \sum_{\Sigma = \text{diag}(\pm 1)} \Sigma Q \Sigma^*$ is diagonal.

Application:

1) Let $H$ be a $m \times n$ matrix with independent entries such that $h_{jk} \sim -h_{jk}$ for all $j,k$.

× Then $H\Sigma \sim H$ $\forall \Sigma$ (ex: $H \begin{pmatrix} -1 & & 0 \\ & 1 & \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} -h_{11} & h_{12} & \cdots & h_{1n} \\ \vdots & & & \\ -h_{m1} & h_{m2} & \cdots & h_{mn} \end{pmatrix} \sim H$),

so lemma 1 applies, i.e. the optimal $Q$ is diagonal.

2) Let $H$ be a $m \times n$ matrix with i.i.d. entries

× Then $H\Pi \sim H$ $\forall \Pi$ (clear), so lemma 2 applies, i.e. the optimal $Q$ has equal diagonal el. & equal off-diag. el.

1+2) Let $H$ be a $m \times n$ matrix with iid entries such that $h_{jk} \sim -h_{jk}$ $\forall j,k$. Then $H \sim H\Sigma \sim H\Pi$ for any $\Sigma$ & $\Pi$, so the optimal $Q$ is of the

× form $Q = \frac{P}{n} I$.

A particular case: "i.i.d. Rayleigh fading" [Telatar 95]

Let us assume that $H$ is a $m \times n$ matrix with i.i.d. entries distributed as $N_{\mathbb{C}}(0,1)$ r.v.

Remark: this assumption is an assumption of the same flavor as that made by Wigner in physics: rather than trying to model precisely the $n \cdot m$ fading coefficients with a complicated deterministic model, let us consider simply the matrix $H$ as "completely random".

The above distribution falls into the above case (1+2), so the optimal input covariance matrix is $Q = \frac{P}{n} I$ and the capacity is given by $C = \mathbb{E}\left(\log\det\left(I + \frac{P}{n} H H^*\right)\right)$.

Remark: an alternate way to derive the result is to notice that in this very particular case, the distribution of $H$ is unitarily invariant, i.e. that $H \sim HU$ for any $n \times n$ unitary matrix $U$. This implies in particular that $H \sim H\Sigma$ and $H \sim H\Pi$, since both $\Sigma$ and $\Pi$ are unitary matrices.

The capacity may be further written as

$$C = \mathbb{E}\left( \sum_{j=1}^{m} \log\left(1 + \frac{P}{n} \lambda_j\right)\right)$$

with $\lambda_j$ the eigenvalues of the $m \times m$ matrix $HH^*$.

For simplicity, let us consider the case where $\underline{m = n}$:

$$\begin{aligned}
C &= \int_{\mathbb{R}_+^n} d\lambda_1 \ldots d\lambda_n \cdot p(\lambda_1 \ldots \lambda_n) \cdot \left( \sum_{j=1}^{n} \log\left(1 + \frac{P}{n} \lambda_j\right)\right) \\
&= \sum_{j=1}^{n} \int_{\mathbb{R}_+} d\lambda_j \, p(\lambda_j) \log\left(1 + \frac{P}{n} \lambda_j\right) \\
&= n \int_{\mathbb{R}_+} d\lambda \, p(\lambda) \log\left(1 + \frac{P}{n} \lambda\right) \\
\text{(see lecture 4)} &= \sum_{\ell=0}^{n-1} \int_{\mathbb{R}_+} d\lambda \, e^{-\lambda} \underbrace{L_\ell(\lambda)}_{\text{(= Laguerre polynomials)}}{}^2 \log\left(1 + \frac{P}{n} \lambda\right)
\end{aligned}$$

It turns out that this expression is proportional to $n$ as $n$ gets large (see the forthcoming asymptotic analysis).

More generally, $C$ is proportional to $\min(m, n)$ when $m \neq n$.

Third scenario: $H$ is a random matrix

that is fixed once and for all ("slow fading")

Moreover, we assume again that the receiver knows the realizations of $H$, but not the transmitter (situation c). We also assume that $H$ is a $m \times n$ matrix with iid $\sim \mathcal{N}_{\mathbb{C}}(0,1)$ entries.

For a given $X$ with input covariance $Q$ and a given $H$, the mutual information between $X$ and $Y = HX + Z$ is $\log\det(I + HQH^*)$.

There is always a strictly positive probability that this expression is arbitrarily small, so the capacity is zero.

We therefore shift our attention to a new quantity; the outage probability: for a target rate $R$, we define

$$\mathbb{P}_{out}(R) = \min_{Q \geq 0 : \operatorname{Tr} Q \leq P} \underbrace{\mathbb{P}\left(\log\det(I + HQH^*) < R\right)}_{= \text{ cumulative distribution function}}$$

This optimization problem is considerably harder than the preceding; actually, the solution is not known!

NB: $C$ = upper bound on achievable rate

$\mathbb{P}_{out}(R)$ = lower bound on error probability

Since $H \sim HU$ for any $n \times n$ unitary matrix $U$ and the constraint is also unitarily invariant, we have

$$\mathbb{P}_{out}(R) = \min_{Q \text{ diag} \geq 0:\ \mathrm{Tr}\, Q \leq P} \mathbb{P}\left( \log \det (I + HQH^*) < R \right)$$

Conjecture [Telatar 95] [resolved so far only for $m=1$]

The optimal input covariance matrix $Q$ is of the form $Q = \mathrm{diag}(\underbrace{\tfrac{P}{k}, \ldots, \tfrac{P}{k}}_{k \text{ times}}, 0, \ldots, 0)$ [the order is irrelevant here]

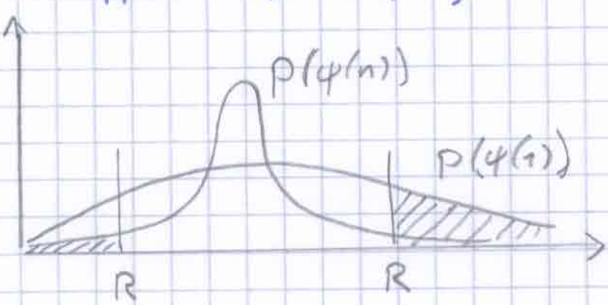
for some $1 \leq k \leq n$.

Remark: the optimal $k$ should depend on the target rate $R$:

Let $\varphi(k) = \log \det \left( I + H \,\mathrm{diag}(\tfrac{P}{k}, \ldots, \tfrac{P}{k}, 0, \ldots, 0) H^* \right)$

= mutual information obtained by using $k$ antennas

- If $R$ is sufficiently small, then $\varphi(n)$ is greater than $R$ with high probability, so $k=n$ is optimal (averaging effect)
- If $R$ is sufficiently large, then $k=1$ is optimal: ("all eggs in one basket")




At high SNR ($P \to \infty$), this problem simplifies and more can be said [next time].