Random matrix theory: lecture 18

Capacity scaling of multi-antenna channels (MIMO)

× • n transmit antennas and n receive antennas (simplifying assumption)

• channel model: $Y = HX + Z$

• Z = additive white Gaussian noise (of variance 1)

• H = channel matrix with iid entries

such that $\mathbb{E}(|h_{jk}|^2) = 1$ & $h_{jk} \sim -h_{jk}$

(NB: implies that $\mathbb{E}(h_{jk}) = 0$)

• global power constraint at the transmitter: $\sum_{j=1}^{n} \mathbb{E}(|x_j|^2) \leq P$

• assumption: only the receiver knows the channel realizations

$$\Rightarrow \quad C_n = \max_{Q \geq 0 : \operatorname{Tr} Q \leq P} \mathbb{E}\left(\log\det\left(I + HQH^*\right)\right)$$

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

(see lecture 6 and lemmas 1 & 2 in lecture 7)

• we are now interested in the scaling of the capacity as $n \to \infty$; let $\lambda_1^{(n)} \ldots \lambda_n^{(n)}$ be the eigenvalues of $W^{(n)} = \frac{1}{n} HH^*$. Then

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

- By the Marčenko-Pastur theorem, we know that $F_n(t) := \frac{1}{n} \#\{j : \lambda_j^{(n)} \leq t\} \xrightarrow[n\to\infty]{} \int_0^t p_\nu(x)\,dx$ a.s. $\forall t \geq 0$

× where $p_\nu(x) = \frac{1}{\pi}\sqrt{\frac{1}{x} - \frac{1}{4}}\; 1_{0<x<4}$ "quater-circle" distribution

- This implies that for any $f : \mathbb{R} \to \mathbb{R}$ bounded and continuous,

$$\frac{1}{n}\sum_{j=1}^{n} f(\lambda_j^{(n)}) \xrightarrow[n\to\infty]{} \int_0^4 f(x)\, p_\nu(x)\,dx \quad \text{a.s.}$$

- Moreover, convergence in expectation also holds:

$$\frac{1}{n}\mathbb{E}\Big(\sum_{j=1}^{n} f(\lambda_j^{(n)})\Big) \xrightarrow[n\to\infty]{} \int_0^4 f(x)\, p_\nu(x)\,dx$$

× - A little extra work is required to show that the theorem holds for $f(x) = \log(1+Px)$, which is continuous on $[0,\infty)$ (note that $\lambda_j^{(n)} \geq 0 \;\forall n$), but unbounded. Therefore, $\frac{1}{n} C_n \xrightarrow[n\to\infty]{} \int_0^4 \log(1+Px)\, p_\nu(x)\,dx$.

× Main conclusion:

The capacity scales linearly with the number of antennas, for a fixed global power budget P.

Note however that the model of i.i.d. channel gains eventually breaks down as $n \to \infty$.

Capacity scaling of ad hoc wireless networks

The previous result relies on the study of the global regime of random matrices. Various extensions of the above result exist for more general channel matrix models. Most of them again rely on the study of the global regime. In the following, we are going to see an example where the largest eigenvalue of random matrices plays a major role.

Model: • $2n$ nodes, distributed independently and uniformly in a square domain of area n. ($\Rightarrow$ constant density of nodes as n increases)

- nodes are paired up at random so as to form n source-destination pairs (logical links)
- individual power constraint P at each node
- attenuation of signals over distance r: $\frac{e^{i\phi}}{r^{\alpha/2}}$ where ϕ is a random phase and $\alpha \geq 2$.
- additive white Gaussian noise at each receiver

Assume now that the n S-D pairs wish to establish communication at a common rate R_n, without the help of any fixed infrastructure, but using other nodes as relays for their communication.

Question: how does the maximum achievable rate R_n scale with n? Likewise, how does the overall capacity of the network $C_n = nR_n$ scale with n?

Previous answers:

- $C_n \cong \sqrt{n}$ is achievable with multi-hop strategy (Gupta-Kumar 2000)
- for $\alpha > 4$, $\sqrt{n}$ is the best we can do (Kumar-Xie 2006)

× New result: (Ozgur - Leveque - Tse 07)

- for $2 \leq \alpha \leq 3$, $C_n = O(n^{2-\alpha/2+\varepsilon})$ $\forall \varepsilon > 0$
- for $\alpha > 3$, $C_n = O(n^{\frac{1}{2}+\varepsilon})$ $\forall \varepsilon > 0$

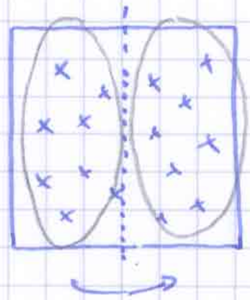
and the first bound is achievable (up to a n^{ε})

× via MIMO and hierarchical cooperation;

the second is achievable via multihop.

Proof of the two upper bounds (simultaneously)

- Cut-set argument (∗) : let us divide the network into two equal parts:



we assume here full cooperation on both sides (turning the network into a MIMO channel)

What is the maximum information flow going from left to right? Since there are statistically order n S-D pairs willing to establish communication from the left-hand side to the right-hand side, this information flow is an upper bound on $n R_n$, i.e. is an upper bound on C_n (up to some constant).

From left to right, we have the following MIMO channel:

$$y_j = \sum_{k=1}^{n} h_{jk} \, x_k + z_j$$

(NB: questionable assumptions)

× where $h_{jk} = \dfrac{e^{i \phi_{jk}}}{r_{jk}^{\alpha/2}}$, with ϕ_{jk} = i.i.d. random phases varying ergodically over time and r_{jk} = distances (fixed).

(∗) = the only tool at hand for dealing with information theoretic capacity of large networks.

At this point we have to specify who knows the channel realizations h_{jk} (i.e. the phases ϕ_{jk}, since we already assume that the positions of the nodes are fixed and known to everybody).

We will study two cases:

1°) only the receivers know the phases ϕ_{jk} (not that realistic in a wireless network, since receivers can always feed these phases back to transmitters)

2°) the phases ϕ_{jk} are known to everybody.

<u>NB</u>: the case where the phases are not known to anybody remains an interesting open problem!

1°) By the above cut-set argument, we have

$$C_n \leq \max_{Q \geq 0 \,:\, Q_{kk} \leq P, \, \forall k} \mathbb{E}\left(\log\det\left(I + H Q H^*\right)\right)$$

Since the h_{jk} are independent and $h_{jk} \sim -h_{jk}$, we have by Lemma 1 of lecture 7 that the optimal covariance matrix Q is diagonal.

But because of the individual power constraint, it is then clear that the optimal $Q = P\,I$, so that

$$C_n \le \mathbb{E}\left(\log\det(I + PHH^*)\right)$$
$$\le \mathbb{E}\left(\mathrm{Tr}(PHH^*)\right)$$

$$\left[\begin{array}{l} \text{since } \underset{(A \geq 0)}{\log\det(I+A)} = \sum_{j=1}^{n} \log(1+\lambda_j^A) \underset{(\log(1+x) \le x)}{\le} \sum_{j=1}^{n} \lambda_j^A = \mathrm{Tr}\,A \end{array}\right]$$

× Note moreover that the last inequality is reasonably tight, since the eigenvalues of HH^* are relatively small.

Therefore,

$$C_n \le P\,\mathbb{E}\left(\sum_{j,k=1}^{n} |h_{jk}|^2\right)$$
$$= P \sum_{j,k=1}^{n} \frac{1}{r_{jk}^{\alpha}} \qquad \left(\text{since } |e^{i\phi_{jk}}| = 1\right)$$

and estimating the above sum leads to

$$\times \qquad \underset{(\triangle\text{ technical detail skipped})}{\sum_{j,k=1}^{n} \frac{1}{r_{jk}^{\alpha}}} \sim \begin{cases} n^{2-\alpha/2} & \text{if } 2 \le \alpha \le 3 \\ \sqrt{n} & \text{if } \alpha > 3 \end{cases} \qquad \#$$

Interpretation:

- if $\alpha \le 3$, long range MIMO communications are worth it, but still limited by the power transfer.
- if $\alpha > 3$, then multi-hop communication is optimal (long range communications are too onerous)

2°) If phases are known to everybody, then we only have the a priori looser upper bound:

$$C_n \leq \mathbb{E}\left(\max_{Q \geq 0 : Q_{kk} \leq P, \forall k} \log\det(I + HQH^*) \right)$$

since the transmitters can theoretically tune their covariance matrix to the channel realization H.

But optimizing Q for a given H is a difficult optimization problem (because of the individual power constraint, which is not unitarily invariant).

Nevertheless, we are only interested in scaling laws here; we are therefore going to show that up to some $\log n$'s, the same upper bound as before applies:

$$C_n \leq \mathbb{E}\left(\max_{Q \geq 0 : Q_{kk} \leq P, \forall k} \mathrm{Tr}(HQH^*) \right)$$

Notice that $\mathrm{Tr}(HQH^*) \leq |||H|||_2^2 \cdot \mathrm{Tr}\, Q$, where $|||H|||_2$ is the spectral norm of H.

$$\left[\begin{array}{l} \text{since } \mathrm{Tr}(HQH^*) = \mathbb{E}_x(\mathrm{Tr}(Hxx^*H^*)) = \mathbb{E}_x(x^*H^*Hx) \\ \quad = \mathbb{E}_x(\|Hx\|_2^2) \leq |||H|||_2^2 \cdot \mathbb{E}_x(\|x\|_2^2) = |||H|||_2^2 \cdot \mathrm{Tr}\, Q \end{array}\right]$$

However since $\mathrm{Tr}\, Q \leq nP$, this upper bound only gives $C_n \leq \mathbb{E}\left(|||H|||_2^2 \right) \cdot nP$
and therefore fails to give the correct order (since it can easily be shown that $\mathbb{E}\left(|||H|||_2^2 \right) \geq cst$).

Instead, let us consider the rescaled matrices

$$\begin{cases} \tilde{h}_{jk} := h_{jk} / \sqrt{d_k} \\ \tilde{Q}_{jk} := \sqrt{d_j}\, Q_{jk} \sqrt{d_k} \end{cases} \quad \text{where } d_k := \sum_{j=1}^{n} \frac{1}{r_{jk}^{\alpha}}$$

The upper bound now becomes:

$$C_n \leq \mathbb{E}\left(\max_{\tilde{Q} \geq 0:\ \tilde{Q}_{kk} \leq P d_k} \mathrm{Tr}(\tilde{H} \tilde{Q} \tilde{H}^*) \right)$$

Using again the inequality $\mathrm{Tr}(\tilde{H}\tilde{Q}\tilde{H}^*) \leq |||\tilde{H}|||_2^2\, \mathrm{Tr}\,\tilde{Q}$, we obtain

$$C_n \leq \mathbb{E}\left(|||\tilde{H}|||_2^2 \right) \cdot \underbrace{P \underbrace{\sum_{k=1}^{n} d_k}_{= \sum_{j,k=1}^{n} \frac{1}{r_{jk}^{\alpha}} = \begin{cases} O(n^{2-\alpha/2}) & 2 \leq \alpha \leq 3 \\ O(\sqrt{n}) & \alpha > 3 \end{cases}}}_{= \text{correct order!}}$$

There remains to prove that this term does not scale faster than some log n's.

Note that $|||\tilde{H}|||_2^2 = \rho(\tilde{H}\tilde{H}^*) = \lim_{\ell\to\infty} \|(\tilde{H}\tilde{H}^*)^\ell\|^{1/\ell}$

for any matrix norm $\|\cdot\|$. Choosing $\|A\| = \sqrt{\mathrm{Tr}(A^*A)}$,

we obtain $|||\tilde{H}|||_2^2 = \lim_{\ell\to\infty} \mathrm{Tr}\left((\tilde{H}\tilde{H}^*)^{2\ell}\right)^{1/2\ell}$

$= \lim_{\ell\to\infty} \mathrm{Tr}\left((\tilde{H}\tilde{H}^*)^\ell\right)^{1/\ell}$. Therefore,

$$\mathbb{E}\left(|||\tilde{H}|||_2^2\right) \underset{DCT}{=} \lim_{\ell\to\infty} \mathbb{E}\left(\mathrm{Tr}\left((\tilde{H}\tilde{H}^*)^\ell\right)^{1/\ell}\right) \underset{\text{Jensen}}{\leq} \lim_{\ell\to\infty} \left(\mathbb{E}\left(\mathrm{Tr}\left((\tilde{H}\tilde{H}^*)^\ell\right)\right)\right)^{1/\ell}$$

The moment computation gives

$$\mathbb{E}\left(\mathrm{Tr}\left((\tilde{H}\tilde{H}^*)^\ell\right)\right) \leq t_\ell\, n (\log n)^{3\ell} \quad (*)$$

so

$$\mathbb{E}\left(|||\tilde{H}|||_2^2\right) \leq \lim_{\ell\to\infty} \left(t_\ell^{1/\ell}\right)(\log n)^3 = 4(\log n)^3 \quad \#$$

(*) illustration for $\ell = 2$ and a 1D regular network:

$$r_{jk} = j+k \implies d_k = \sum_{j=1}^{n} \frac{1}{(j+k)^\alpha} \sim k^{1-\alpha}$$

$$\implies \tilde{h}_{jk} \cong \frac{e^{i\phi_{jk}}}{(j+k)^{\alpha/2}} \sqrt{k^{\alpha-1}} \; : \; |\tilde{h}_{jk}|^2 = \frac{k^{\alpha-1}}{(j+k)^\alpha} \leq \frac{1}{j+k}$$

$$\mathbb{E}\left(\mathrm{Tr}\left((\tilde{H}\tilde{H}^*)^2\right)\right) = \sum_{j_1,j_2,k_1,k_2=1}^{n} \mathbb{E}\left(\tilde{h}_{j_1k_1}\overline{\tilde{h}_{j_2k_1}}\tilde{h}_{j_2k_2}\overline{\tilde{h}_{j_1k_2}}\right)$$

$$= \sum_{j,k_1\neq k_2} \mathbb{E}(|\tilde{h}_{jk_1}|^2)\,\mathbb{E}(|\tilde{h}_{jk_2}|^2) + \sum_{j_1\neq j_2,k} \mathbb{E}(|\tilde{h}_{j_1k}|^2)\,\mathbb{E}(|\tilde{h}_{j_2k}|^2)$$

$$+ \sum_{j,k} \mathbb{E}(|\tilde{h}_{jk}|^4) \leq \sum_{j,k_1\neq k_2} \frac{1}{j+k_1}\frac{1}{j+k_2} + \sum_{j_1\neq j_2,k} \frac{1}{j_1+k}\frac{1}{j_2+k} + \sum_{j,k}\frac{1}{(j+k)^2}$$

$$\leq 2\sum_j \Big(\underbrace{\sum_k \frac{1}{j+k}}_{\leq \log n}\Big)^2 \leq \underbrace{2}_{=t_2}\, n \underbrace{(\log n)^2}_{\text{or } (\log n)^6 \text{ for random placement of nodes}}$$

(binning argument)