Capacity scaling of multi-antenna channels (MIMO)

- $n$ transmit antennas and $n$ receive antennas (assumption)
- Channel model: $Y = HX + Z$
- $Z =$ additive white Gaussian noise (of variance 1)
- $H =$ channel matrix with iid entries such that $\mathbb{E}(|h_{jk}|^2) = 1$ & $h_{jk} = h_{jk}$
  \[ (\text{NB: implies that } \mathbb{E}(h_{jk}) = 0) \]
- Global power constraint at the transmitter: $\sum_{j=1}^{n} \mathbb{E}(|y_j|^2) \leq P$
- Assumption: only the receiver knows the channel realizations

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

$= \mathbb{E}(\log \det (I + P/nHH^*))$

(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 $W^{(n)} = 1/nHH^*$. Then

$C_n = \mathbb{E}\left( \sum_{j=1}^{n} \log \left( 1 + P \lambda_j^{(n)} \right) \right)$
• By the Marcenko-Pastur Theorem, we know that
  \[ F_n(t) := \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{a_j^{(m)} \leq t} \xrightarrow{n \to \infty} \int_0^t p_u(x) \, dx \text{ as } n \to \infty, \]
  where \( p_u(x) = \frac{1}{2\sqrt{x}} \mathbb{1}_{0 \leq x \leq 4} \) is called the "quarter-circle" distribution.

• This implies that for any \( f : \mathbb{R} \to \mathbb{R} \) bounded and continuous,
  \[ \frac{1}{n} \sum_{j=1}^{n} f(a_j^{(m)}) \xrightarrow{n \to \infty} \int_0^4 f(x) p_u(x) \, dx \text{ a.s.} \]

• Moreover, convergence in expectation also holds:
  \[ \frac{1}{n} \mathbb{E} \left( \sum_{j=1}^{n} f(a_j^{(m)}) \right) \xrightarrow{n \to \infty} \int_0^4 f(x) p_u(x) \, dx \]

• 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 \( a_j^{(m)} \geq 0 \) \( \forall n \)), but unbounded.
  Therefore, \( \frac{1}{n} C_n \xrightarrow{n \to \infty} \int_0^4 \log(1+Px) p_u(x) \, dx \).

---

**Final 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 \text{constant density of nodes as } n \text{ 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 \( \phi \) is a random phase and \( \alpha > 2 \)
- 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:**
- $C_n = \sqrt{n}$ is achievable with multi-hop strategy (Gupta-Kumar 2000)
- for $\alpha > 4$, $\sqrt{n}$ is the best we can do (Kumar-Kie 2005)

**New result:** (Ozgur-Leveque-Tse 07)
- for $2 \leq \alpha \leq 3$, $C_n = O(n^{2-\alpha/2+\varepsilon})$ for all $\varepsilon > 0$
- for $\alpha > 3$, $C_n = O(n^{\frac{\alpha}{2}+\varepsilon})$ for all $\varepsilon > 0$

and the first bound is achievable (up to a $n^2$) 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 \( nR_n \), i.e., is an upper bound on \( C_n \) (up to some constant).

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

\[
g_{j} = \sum_{k=1}^{n} h_{jk} x_k + z_j
\]

\( (N) \) (inescapable assumptions)

Where \( h_{jk} = e^{i \phi_{jk}} \), with \( \phi_{jk} \) iid. 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 $h_{j,k}$ (i.e. the phases $\phi_{j,k}$, 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 $\phi_{j,k}$ (not that realistic in a wireless network, since receivers can always feed these phases back to transmitters)

2) the phases $\phi_{j,k}$ are known to everybody.

---

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

---

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

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

Since the $h_{j,k}$ are independent and $h_{j,k} \sim \mathcal{C}\mathcal{N}(0,1)$, 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

$$C_n \leq E \left( \log \det (I + PHH^*) \right)$$

$$\leq E \left( \text{Tr} \left( PHH^* \right) \right)$$

$$(A \geq 0) \quad \quad (\log(nx) \leq x)$$

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

$$C_n \leq P E \left( \sum_{j,k=1}^{n} |h_{jk}|^2 \right)$$

$$= P \sum_{j,k=1}^{n} \frac{1}{\sigma_{jk}^2} \quad (\text{since } |e^{i\theta_{jk}}| = 1)$$

and estimating the above sum leads to

$$\sum_{j,k=1}^{n} \frac{1}{\sigma_{jk}^2} \sim \begin{cases} \frac{n^{2-\alpha}}{\sqrt{n}} & \text{if } 2 < \alpha \leq 3 \\ \sqrt{n} & \text{if } \alpha > 3 \end{cases}$$

(\text{A technical detail skipped})

**Interpretation:**

- If $\alpha \leq 3$, long range \textit{NLoS} communications are worth it, but still limited by the power transfer.
- If $\alpha > 3$, then multi-hop communications is optimal (long range communications are too onerous).
If phases are known to everybody, then we only have the a priori looser upper bound:

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

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:

$$C_n \leq E \left( \max_{Q \succeq 0 : Q_{kk} \leq P, k} \text{Tr} \left( H Q H^* \right) \right)$$

Notice that $\text{Tr} \left( H Q H^* \right) \leq ||H||_2^2 \cdot \text{Tr} Q$,

where $||H||_2$ is the spectral norm of $H$.

$$\text{since } \text{Tr} \left( H Q H^* \right) = E_x \left( \text{Tr} \left( H x x^* H^* \right) \right) = E_x \left( x^* H^* H x \right)$$

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

Instead, let us consider the rescaled matrices 
\[
\begin{align*}
\tilde{h}_{jk} &:= h_{jk} / \sqrt{d_k} \\
\tilde{q}_{jk} &:= \sqrt{d_j} \cdot q_{jk} / \sqrt{d_k}
\end{align*}
\]

The upper bound now becomes:

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

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

\[
C_n \leq \mathbb{E}\ (\|H\|_2^2) \cdot P \sum_{k=1}^{n} \frac{1}{d_k}
\]

\[
= \sum_{j=1}^{n} \frac{1}{r_{jk}^2} = \begin{cases} O(n^{2-\alpha}) & \text{if } \alpha > 3 \\ O(\ln n) & \text{if } \alpha < 3 \end{cases}
\]

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

= correct order!
Note that \( \| W_1 \|_2^2 = e(\hat{W}_1, \hat{W}_1^*) = \lim_{\ell \to \infty} \ell \| (\hat{W}_1, \hat{W}_1^*) \|_{\ell^2} \)

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

we obtain \( \| W_1 \|_2^2 = \lim_{\ell \to \infty} \text{Tr} \left( (\hat{W}_1, \hat{W}_1^*)^2 \right) \) \( \ell \) \( e \)

\( = \lim_{\ell \to \infty} \text{Tr} \left( (\hat{W}_1, \hat{W}_1^*)^2 \right) \) \( \ell \) \( e \). Therefore,

\[
E(\| W_1 \|_2^2) = \lim_{\ell \to \infty} E \left( \text{Tr} \left( (\hat{W}_1, \hat{W}_1^*)^2 \right) \right) \leq \lim_{\ell \to \infty} \left( E \left( \text{Tr} \left( (\hat{W}_1, \hat{W}_1^*)^2 \right) \right) \right) \ell \) \( e \)

The moment computation gives

\[
E \left( \text{Tr} \left( (\hat{W}_1, \hat{W}_1^*)^2 \right) \right) \leq t_2 n \left( \log n \right)^2 \]

(*)

so \( E(\| W_1 \|_2^2) \leq \lim_{\ell \to \infty} \left( (t_2 \ell) \left( \log n \right)^2 \right) = 4 \left( \log n \right)^3 \)

(*) Illustration for \( l = 2 \) and a 1D regular network:

\[
g_{jk} = j + k \quad \Rightarrow \quad d_k = \sum_{j=1}^{n} \frac{1}{(j+k)^2} \sim k^{-\alpha}
\]

\[
\Rightarrow \quad \hat{g}_{jk} = \frac{e^{i\phi_{jk}}}{(j+k)^\alpha} \sqrt{k^{n-\alpha}} \quad \text{such that} \quad \left| \hat{h}_{jk} \right|^2 = \frac{k^{n-\alpha}}{(j+k)^\alpha} \leq \frac{1}{j+k}
\]

\[
E \left( \text{Tr} \left( (\hat{W}_1, \hat{W}_1^*)^2 \right) \right) = \sum_{j, k, l, h, k_2} \mathbb{E} \left( \hat{g}_{j, l, k_1, k_2, k} \bar{g}_{j, l, k_1, k_2, l} \bar{g}_{j, l, k_2, k_1, k} \right)
\]

\[
= \sum_{j, k, k_1, k_2} \mathbb{E} \left( \left| \hat{h}_{jk} \right|^2 \right) \mathbb{E} \left( \left| \hat{h}_{jk} \right|^2 \right) + \sum_{j, k, k_1, k_2} \mathbb{E} \left( \left| \hat{h}_{j1, k_1} \right|^2 \right) \mathbb{E} \left( \left| \hat{h}_{jk} \right|^2 \right) + \sum_{j, k, k_1, k_2} \mathbb{E} \left( \left| \hat{h}_{j1, k_1} \right|^2 \right) \mathbb{E} \left( \left| \hat{h}_{jk} \right|^2 \right)
\]

\[
\leq 2 \sum_{j, k, l, h, k_1, k_2} \left( \frac{1}{j+k} \right)^2 \leq 2 n \left( \log n \right)^2
\]

\[
\leq \log n = t_2 \quad \text{or} \quad (\log n)^6 \quad \text{for random placement of nodes (binning argument)}
\]